

Requirements for processing data from an AmeriFlux site using ONEFlux

AMP webinar series
October 19th, 2020

What to expect

- Logistics
 - About 45 minutes of content, 1 hour total
 - A few stop points through the presentation for quick questions
 - Longer Q&A at the end
 - Live streaming (and recording/slides available at ameriflux.lbl.gov)
- This *is not*:
 - How to run ONEFlux
 - How to use AmeriFlux / FLUXNET data products
- This *is*:
 - What is needed to create data products with ONEFlux for an AmeriFlux site
- Familiarity with AmeriFlux data upload process helps (not required)

What to expect

- Logistics
 - About 45 minutes of content, 1 hour total
 - A few stop points through the presentation for quick questions
 - Longer Q&A at the end
 - Live streaming (and recording/slides available at ameriflux.lbl.gov)
- This *is not*:
 - How to run ONEFlux
 - How to use AmeriFlux / FLUXNET data products
- This *is*:
 - What is needed to create data products with ONEFlux for an AmeriFlux site
- Familiarity with AmeriFlux data upload process helps (not required)



Danielle
Christianson



Housen
Chu



Gilberto
Pastorello



Dario
Papale

Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products

Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products



Gilberto
Pastorello

FLUXNET Datasets

- Marconi 2000: 38 sites / 97 site-years
 - Data contributed by attendees of workshop
 - Access limited to attendees
- LaThuile 2007: 252 sites / 965 site-years
 - Data contributed by many site PIs
 - Start of uniform code base and processing for comparable data products
 - Data access/usage opened to other users, with some restrictions (three tiers)
 - Over 500 unique users downloaded data
- FLUXNET2015: 212 sites / 1532 site-years
 - Standardized code base (developed in collaboration among networks)
 - Extensive QA/QC of data (data more comparable and usable, but fewer sites)
 - Data access to anyone, usage has fewer restrictions (two tiers) – tools to keep track of usage
 - Over 3000 unique users downloaded data
 - ***NEW (Feb 2020): 206 sites following data policy based on CC-BY-4.0***

Eddy Covariance Data Processing

- High-frequency data

- Alteddy (ALTERRA),
- EddyPro (LI-COR),
- eddy4R (NEON)
- EdiRE (U. Edinburgh),
- Custom codes

- Data products from fluxes

USTAR filtering, met/flux gap-filling, flux partitioning, uncertainty estimates, etc.

- REddyProc (MPI),
- **ONEFlux (AMP/EUDB/ICOS-ETC)**
- Custom codes

Eddy Covariance Data Processing

- High-frequency data

- Alteddy (ALTERRA),
- EddyPro (LI-COR),
- eddy4R (NEON)
- EdiRE (U. Edinburgh),
- Custom codes

- Data products from fluxes

USTAR filtering, met/flux gap-filling, flux partitioning, uncertainty estimates, etc.

- REddyProc (MPI),
- **ONEFlux (AMP/EUDB/ICOS-ETC)**
- Custom codes

Why one more software?

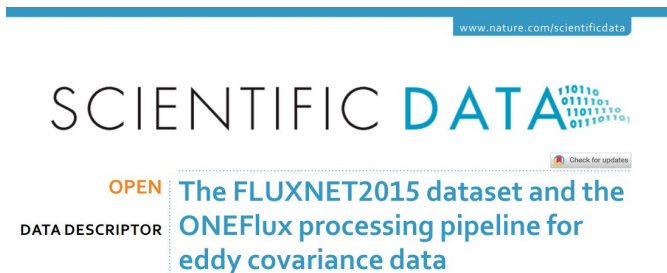
ONEFlux is a package of eddy covariance data processing codes

- *Extensive validation on many site characteristics*
- *Focus on creating network-level data products; standardized for synthesis studies and cross-site comparisons*

ONEFlux

ONEFlux is a package of eddy covariance data processing codes

Reference paper:



Pastorello, G., Trotta, C., Canfora, E. *et al.*
(287 co-authors). ***The FLUXNET2015 dataset
and the ONEFlux processing pipeline for
eddy covariance data.***

Scientific Data **7**, 225 (2020).

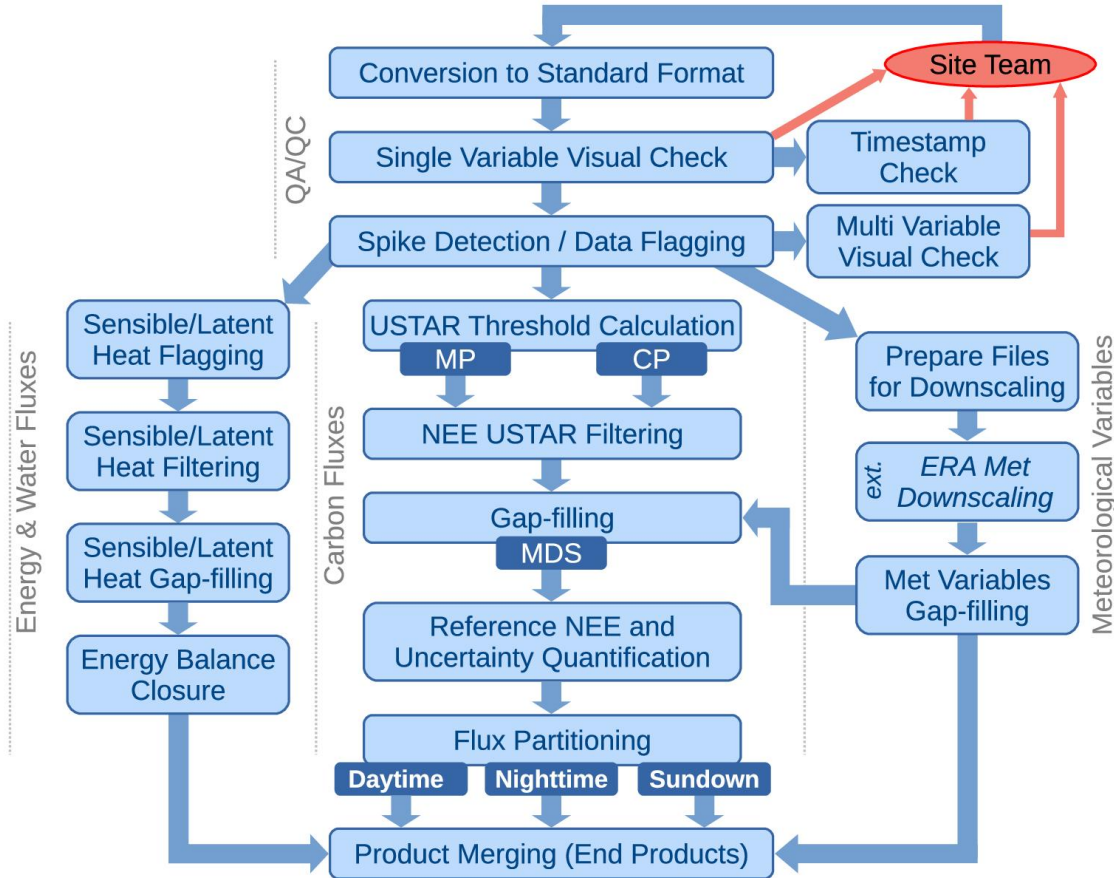
<https://doi.org/10.1038/s41597-020-0534-3>

Code available:

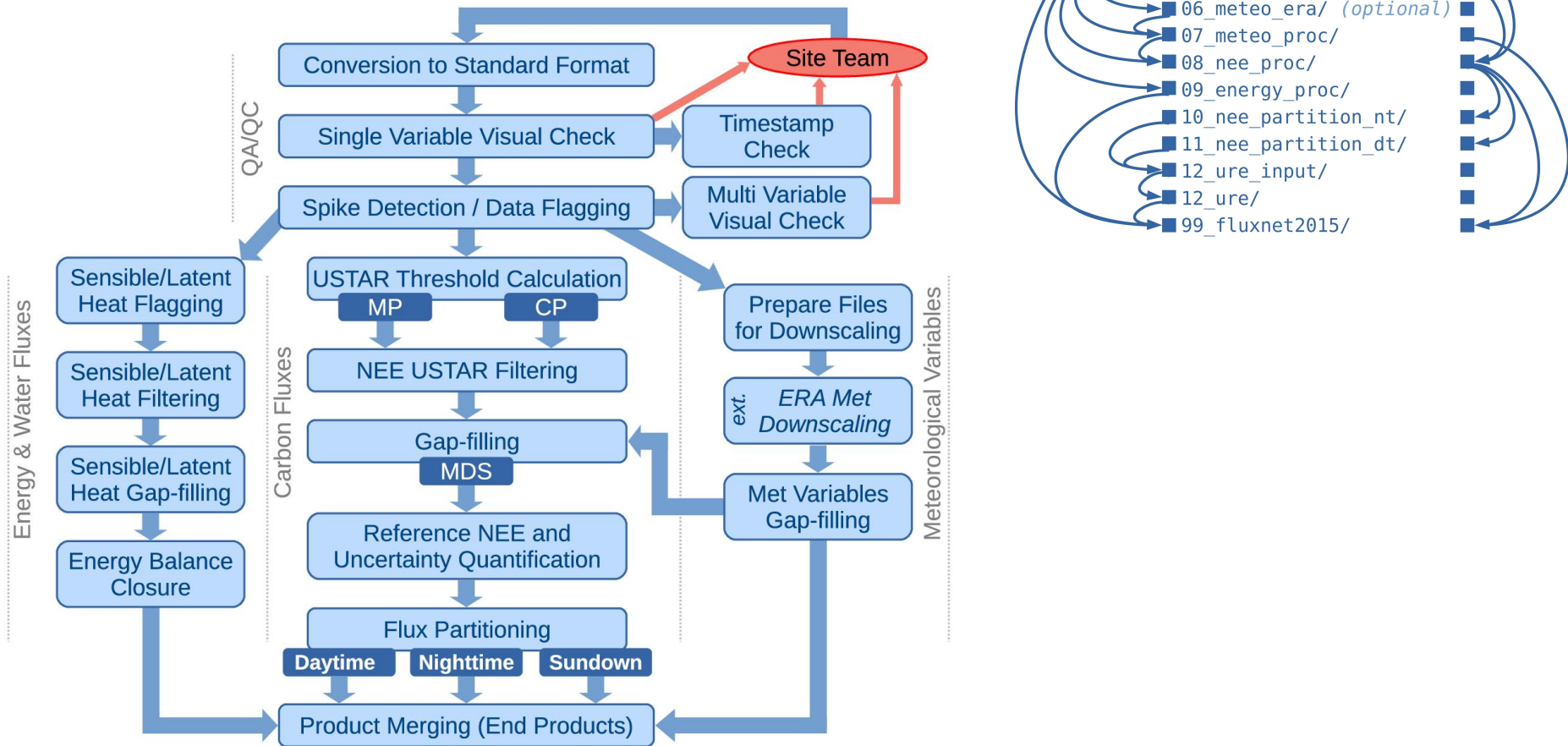


<https://github.com/FLUXNET/ONEFlux>

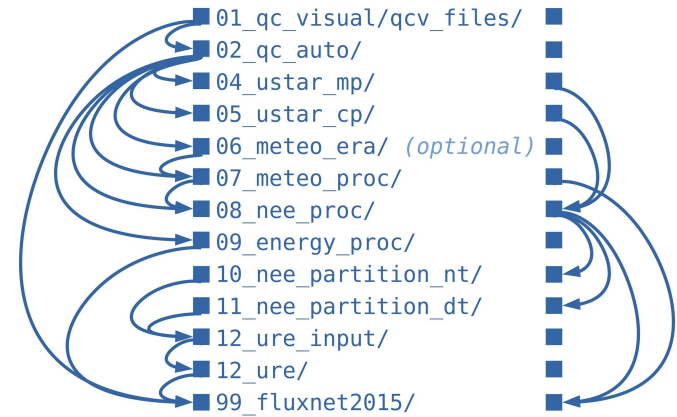
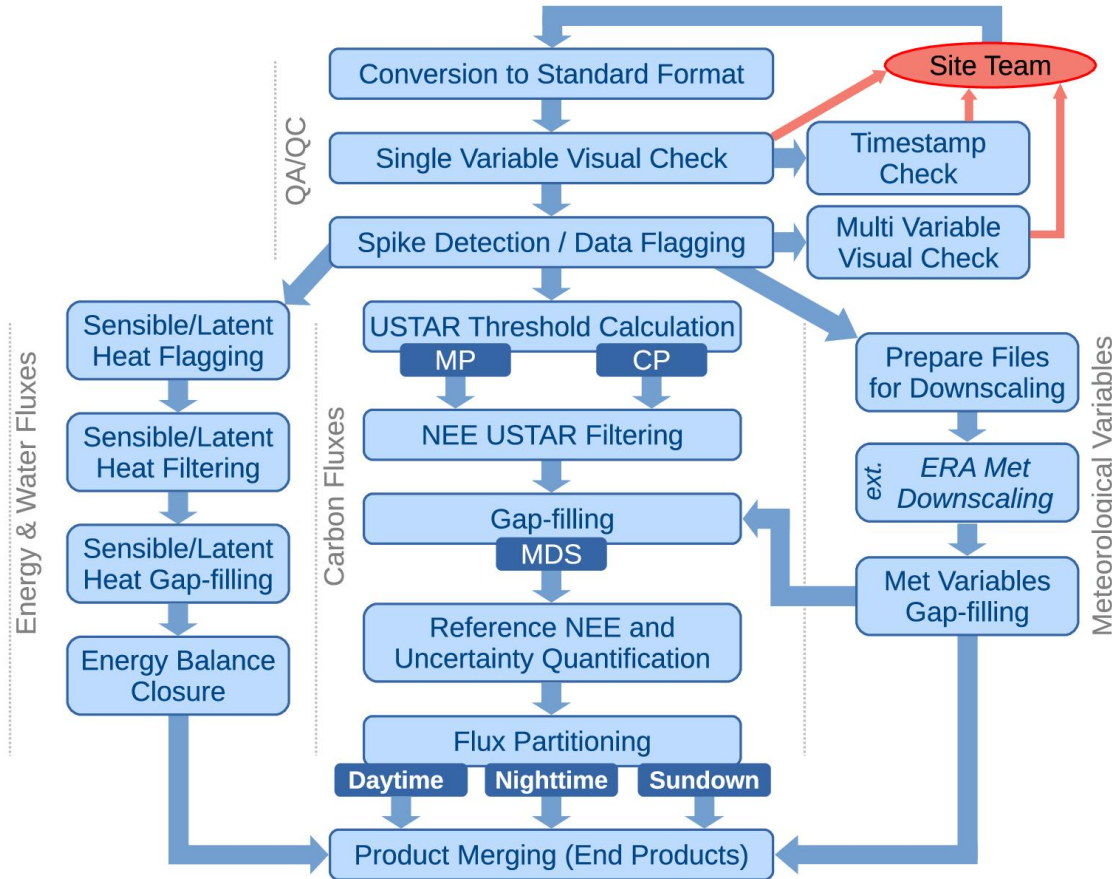
ONEFlux Steps



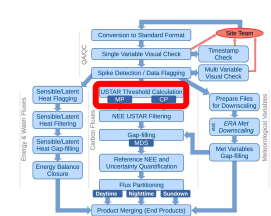
ONEFlux Steps



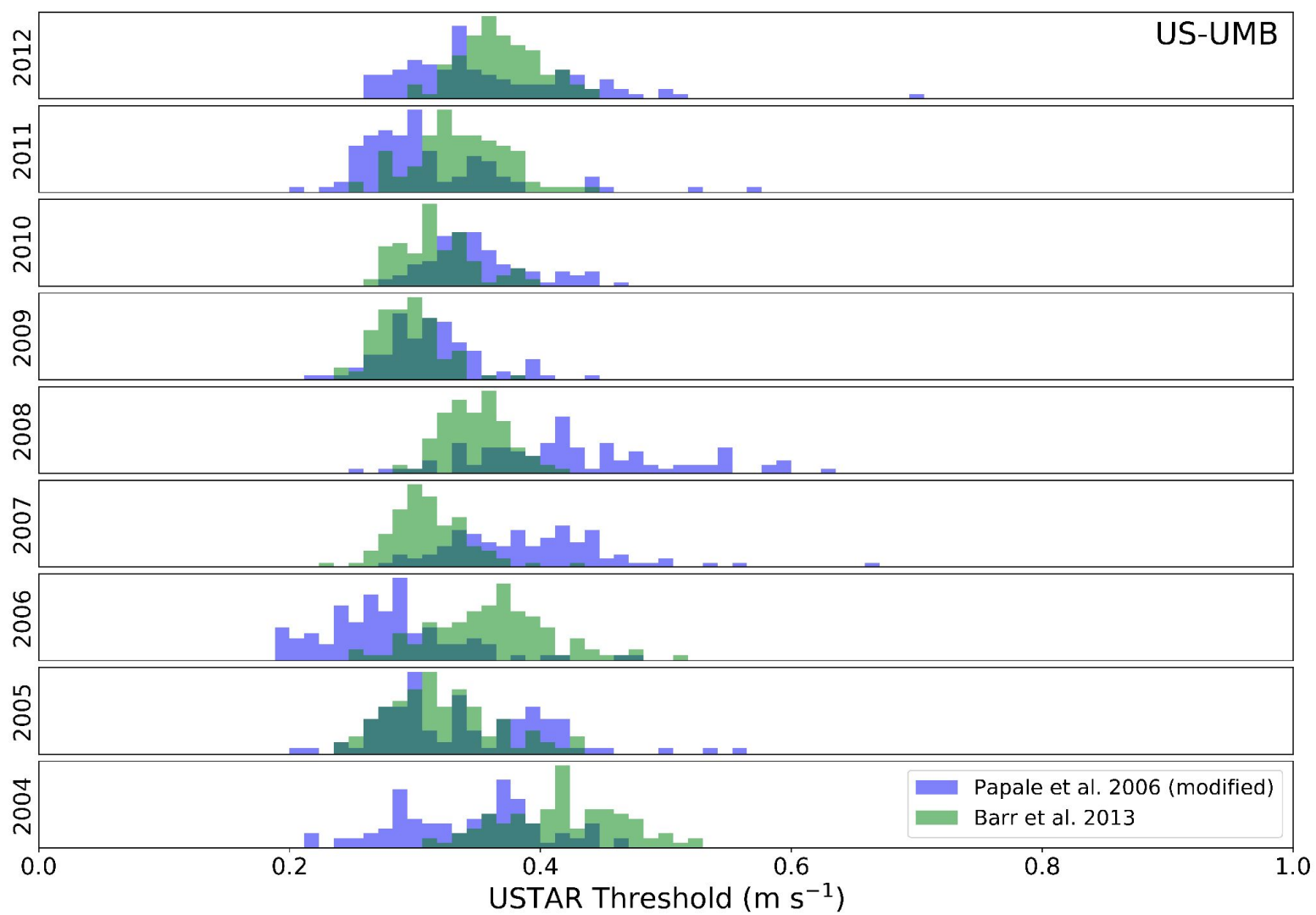
ONEFlux Steps



- 01: input (not-quite FP-Standard)
- 02: flagging/filtering
- 04: USTAR threshold method 1 (unc.)
- 05: USTAR threshold method 2 (unc.)
- 06: downscaling for met gapfilling
- 07: met gapfilling (MDS + downscaling)
- 08: NEE filter/storage/unc./ref/gapfilling
- 09: H & LE corr. factors/gapfilling/unc.
- 10: NEE partitioning (nighttime method)
- 11: NEE partitioning (daytime method)
- 12 (i): create input for 12
- 12: unc. from multiple USTAR thresholds
- 99: merge into FP-Standard compliant output



Distribution of USTAR Threshold values



n.b.

- Key uncertainty estimate
- USTAR should not be pre-filtered

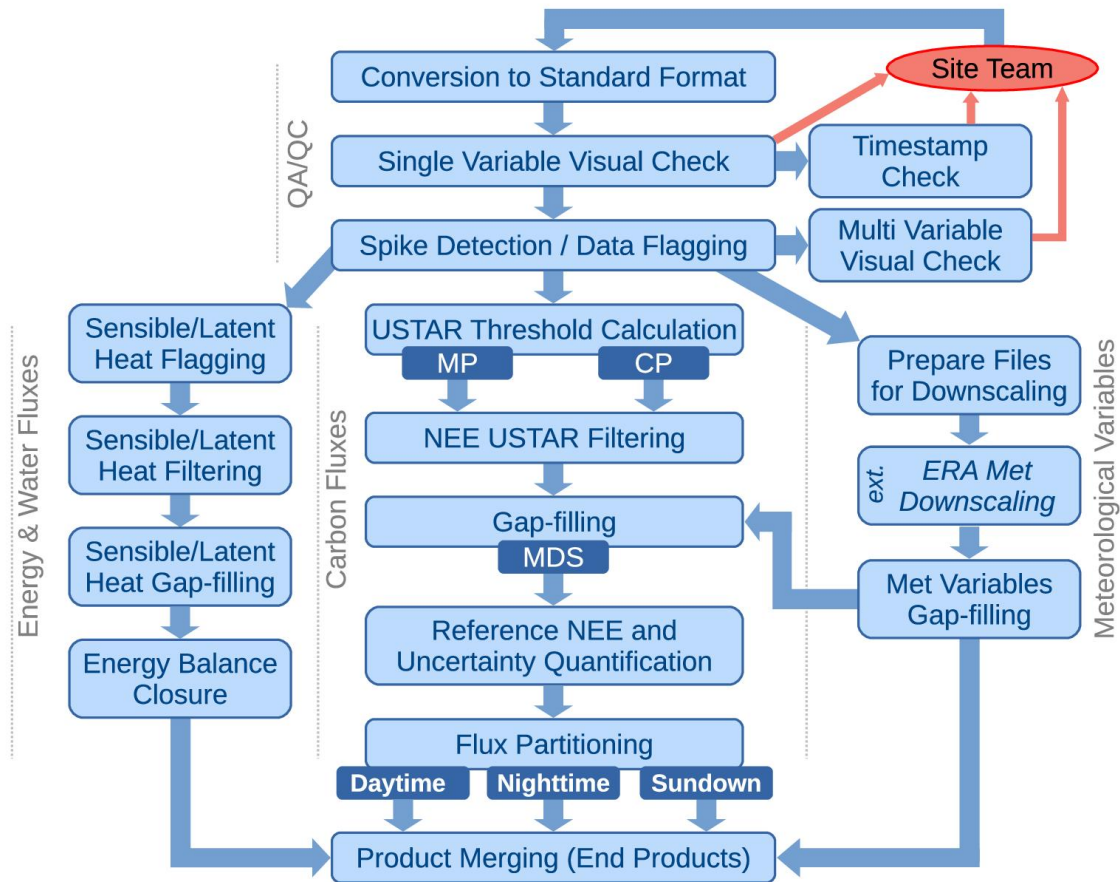
Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products

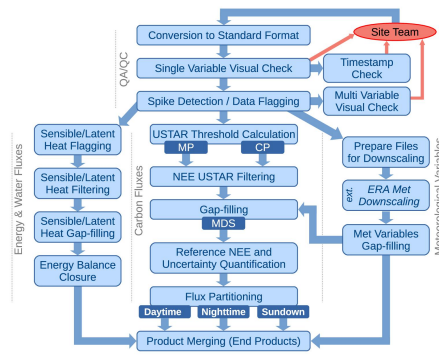


Gilberto
Pastorello

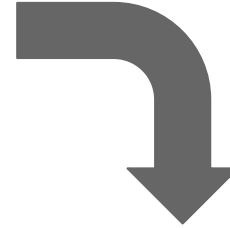
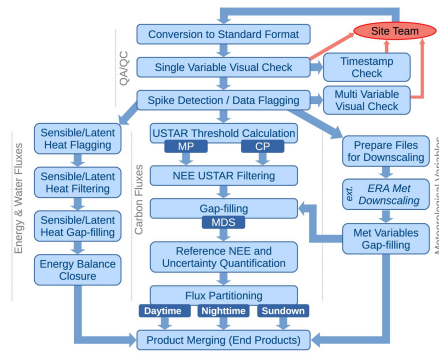
Workflow of a ONEFlux run



Workflow of a ONEFlux run

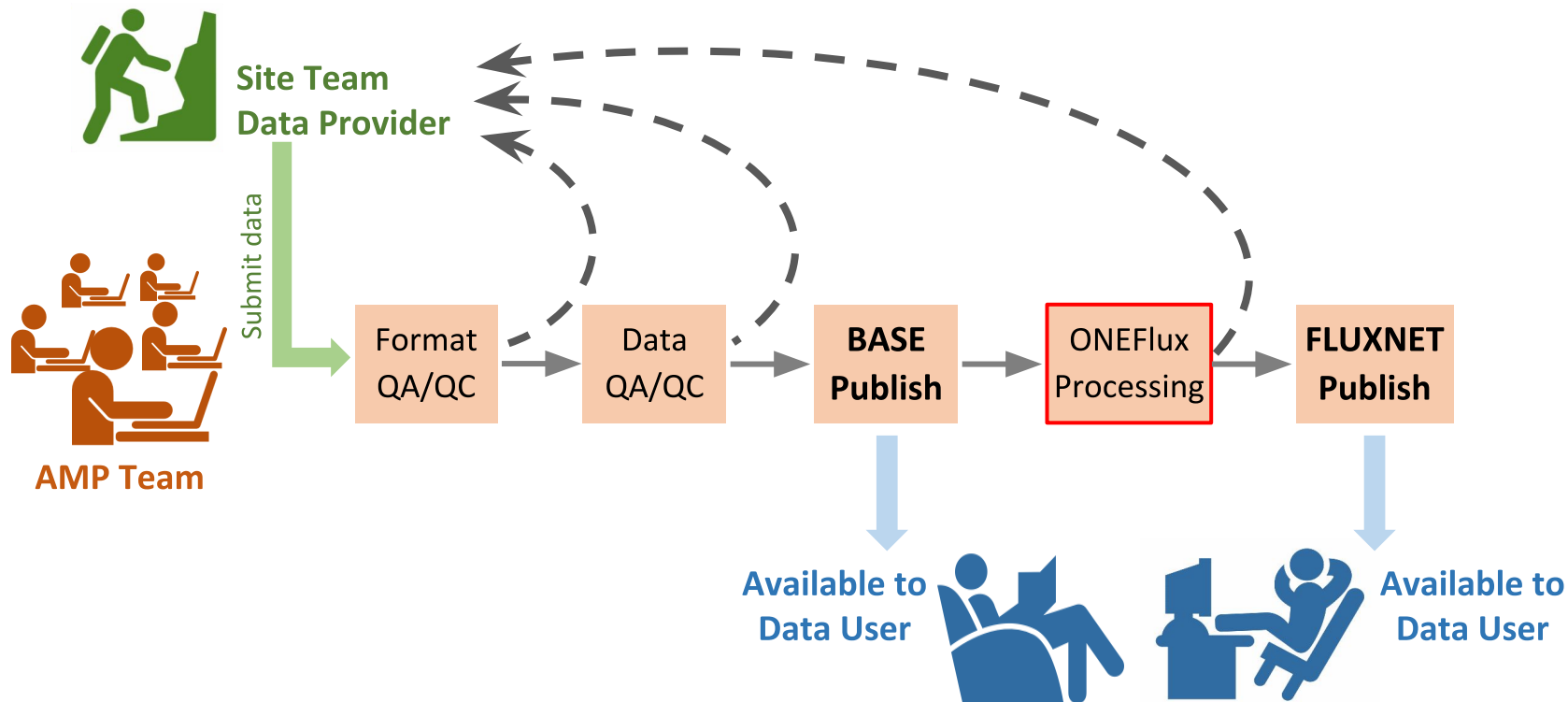


Workflow of a ONEFlux run

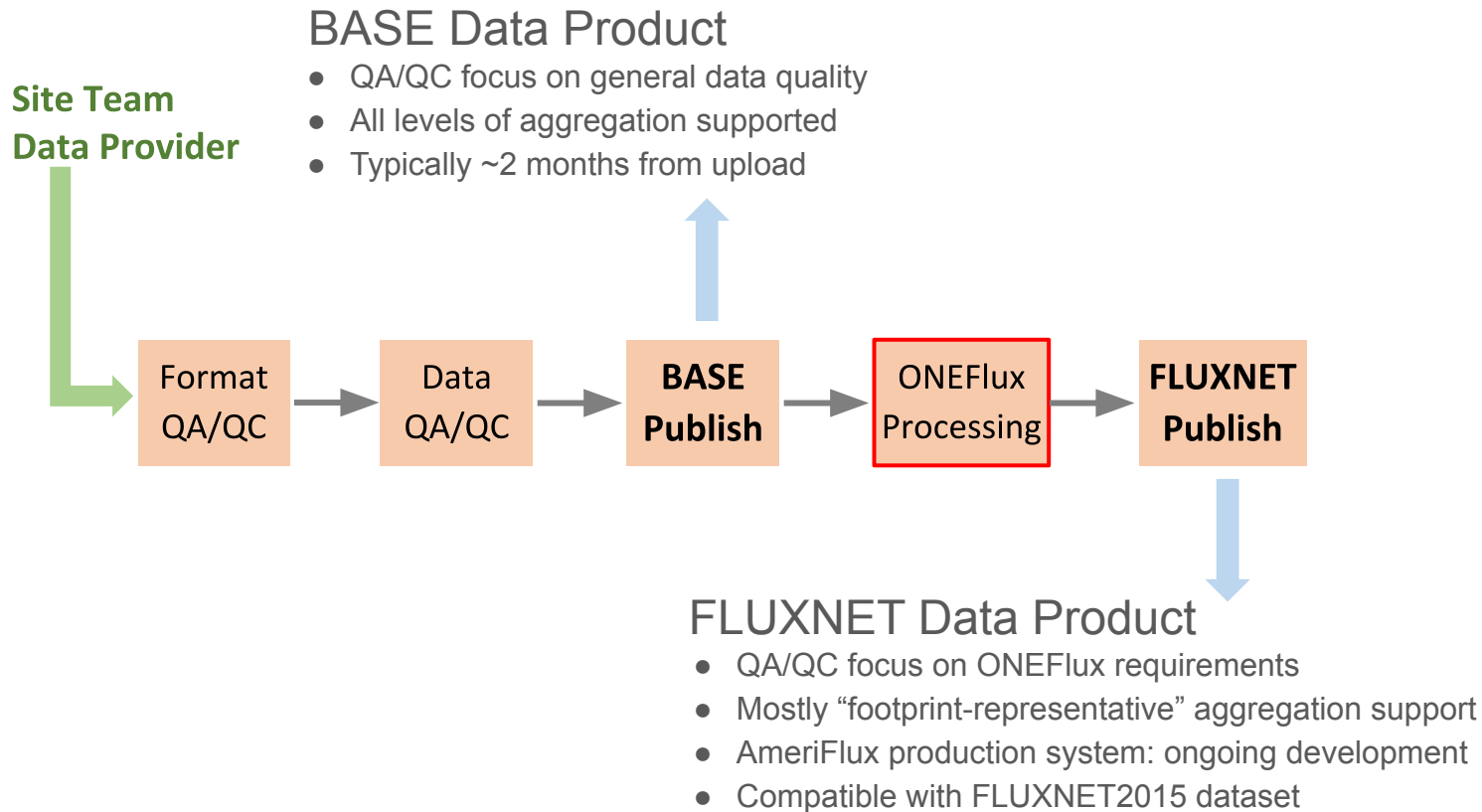


ONEFlux
Processing

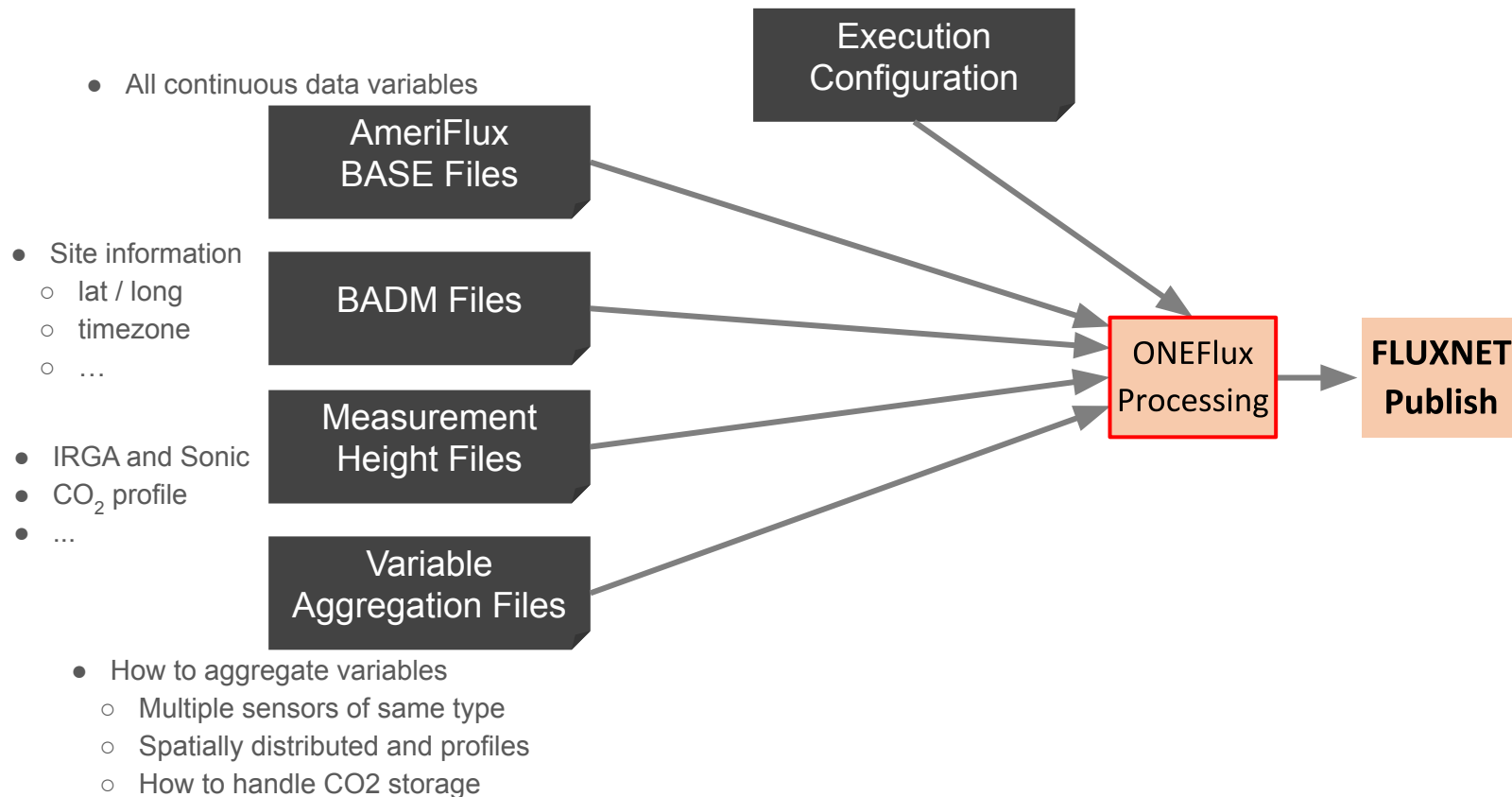
Workflow of a ONEFlux run (within AmeriFlux Data Processing)



Workflow of a ONEFlux run: **Data Products**



Workflow of a ONEFlux run: Required Inputs



Critical data variables for ONEFlux

Required

- **CO₂** ($\mu\text{molCO}_2 \text{ mol}^{-1}$): Carbon Dioxide (CO₂) mole fraction in moist air
- **FC** ($\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$): Carbon Dioxide (CO₂) turbulent flux (without storage component)
- **SC** ($\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$): Carbon Dioxide (CO₂) storage flux measured with a vertical profile system, optional if tower shorter than 3 m
- **H** (W m^{-2}): sensible heat turbulent flux, without storage correction
- **LE** (W m^{-2}): latent heat turbulent flux, without storage correction
- **WS** (m s^{-1}): horizontal wind speed
- **USTAR** (m s^{-1}): friction velocity
- **TA** (deg C): air temperature
- **RH** (%): relative humidity (range 0–100%)
- **PA** (kPa): atmospheric pressure
- **SW_IN** (W m^{-2}): incoming shortwave radiation

Recommended

- **G** (W m^{-2}): ground heat flux, not mandatory, but needed for the energy balance closure calculations
- **NETRAD** (W m^{-2}): net radiation, not mandatory, but needed for the energy balance closure calculations
- **SW_IN_POT** (W m^{-2}): potential incoming shortwave radiation (top of atmosphere theoretical maximum radiation), calculated based on the site coordinates

Suggested

- **PPFD_IN** ($\mu\text{molPhotons m}^{-2} \text{ s}^{-1}$): incoming photosynthetic photon flux density
- **P** (mm): precipitation total of each 30 or 60 minute period
- **LW_IN** (W m^{-2}): incoming (down-welling) longwave radiation
- **SWC** (%): soil water content (volumetric), range 0–100%
- **TS** (deg C): soil temperature

Critical data variables for ONEFlux

Required

- **CO₂** ($\mu\text{molCO}_2 \text{ mol}^{-1}$): Carbon Dioxide (CO₂) mole fraction in moist air
- **FC** ($\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$): Carbon Dioxide (CO₂) turbulent flux (without storage component)
- **SC** ($\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$): Carbon Dioxide (CO₂) storage flux measured with a vertical profile system, optional if tower shorter than 3 m
- **H** (W m^{-2}): sensible heat turbulent flux, without storage correction
- **LE** (W m^{-2}): latent heat turbulent flux, without storage correction
- **WS** (m s^{-1}): horizontal wind speed
- **USTAR** (m s^{-1}): friction velocity
- **TA** (deg C): air temperature
- **RH** (%): relative humidity (range 0–100%)
- **PA** (kPa): atmospheric pressure
- **SW_IN** (W m^{-2}): incoming shortwave radiation

} VPD computed from TA and RH; if provided, used for validation in QA/QC

Recommended

- **G** (W m^{-2}): ground heat flux, not mandatory, but needed for the energy balance closure calculations
- **NETRAD** (W m^{-2}): net radiation, not mandatory, but needed for the energy balance closure calculations
- **SW_IN_POT** (W m^{-2}): potential incoming shortwave radiation (top of atmosphere theoretical maximum radiation), calculated based on the site coordinates

Suggested

- **PPFD_IN** ($\mu\text{molPhotons m}^{-2} \text{ s}^{-1}$): incoming photosynthetic photon flux density
- **P** (mm): precipitation total of each 30 or 60 minute period
- **LW_IN** (W m^{-2}): incoming (down-welling) longwave radiation
- **SWC** (%): soil water content (volumetric), range 0–100%
- **TS** (deg C): soil temperature

} Support for multiple layers of soil variables (SWC & TS)

Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- **Metadata collection for ONEFlux processing**
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products



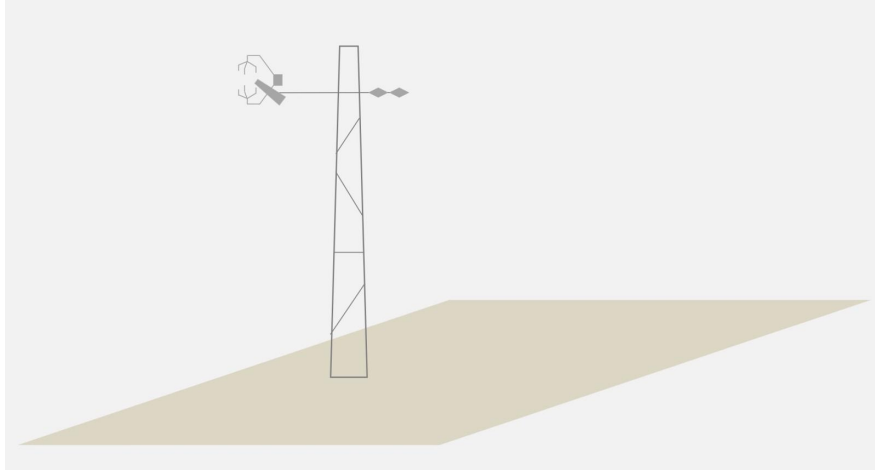
Danielle
Christianson

Metadata collection for ONEFlux processing

BADM = Biological, Ancillary, Disturbance, and Metadata

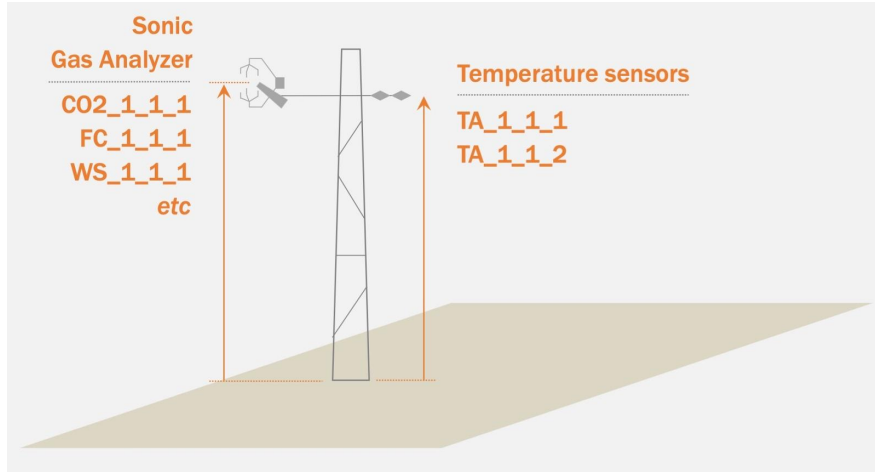
- Variable Information
- Variable Aggregation (and Representation)

Metadata collection for flux-met data variables



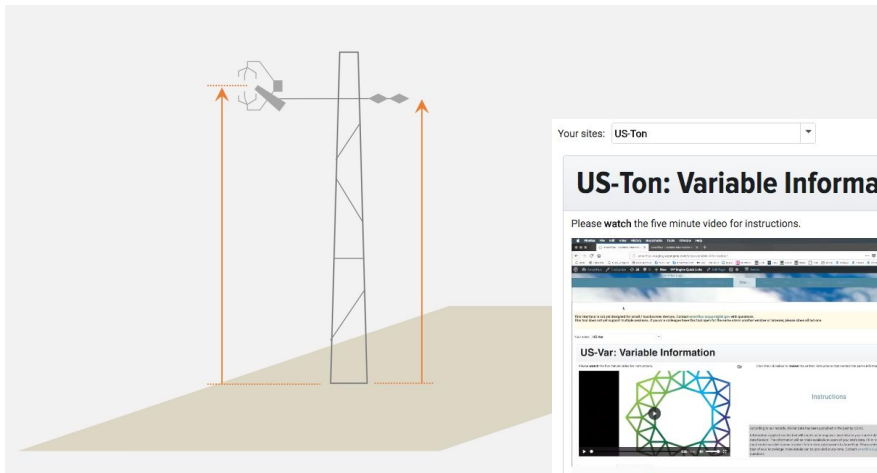
Metadata collection for flux-met data variables

<https://ameriflux.lbl.gov/data/data-variable-qualifier-examples/>



Variable Information: Height and Sensor Info

<https://ameriflux.lbl.gov/sites/variable-information/>



Your sites: US-Ton

US-Ton: Variable Information

Please watch the five minute video for instructions.

Or [Click the link below to review the written instructions that contain the same information as the video:](#)

[Instructions](#)

According to our records, US-Ton data have been published in the past by CDIAC.

Information supplied via this tool will enable us to map your past data to your current data and your current data forward. The information will be made available to users of your site's data. FP-In Variables entered here must match variable names in prior / future data submissions to AmeriFlux. Please enter information to the best of your knowledge; more details can be provided at any time. Contact ameriflux-support@lbl.gov with questions.

Steps

1. Compile or update a list of instrument models used at your site on the *Instrument Model List* tab ([instructions](#)).
2. Verify that previously submitted data variables (gray columns) are mapped to the correct FP-In Variable on the *Variable Information* tab ([instructions](#)).
3. Associate data with height and instrument model information on the *Variable Information* tab ([instructions](#)).

Tips

- Click on throughout the page for additional instructions.
- See FP-In format details at [Half-Hourly / Hourly Data Upload Format and Data Variables](#).
- For special cases, see the [FAQ](#).

Instrument Model List **Variable Information**

[Save](#)

[+ New Variable](#)

Gray columns are data in older formats and cannot be edited. Hover over column for details. View only useful columns with [Customize Columns](#) at right.

Status

☒ Reviewed (Saved)

☐ Updated (NOT Saved)

☐ Not yet addressed (Saved)

☐ To be deleted (⌫) to undo)

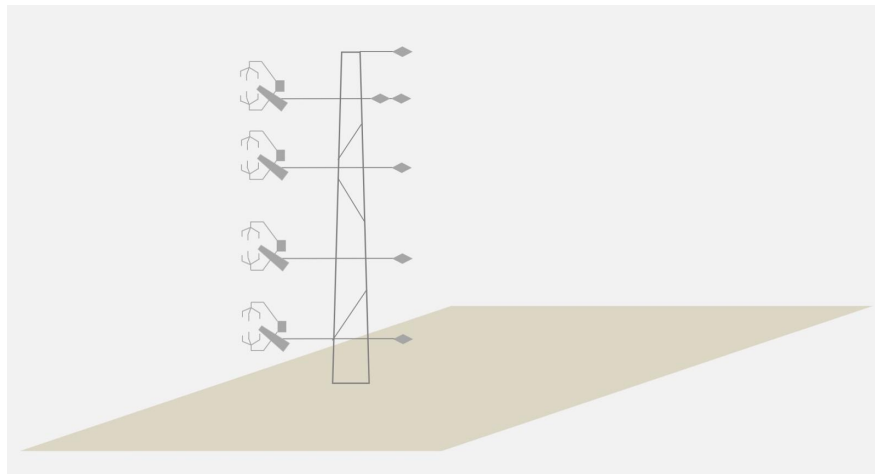
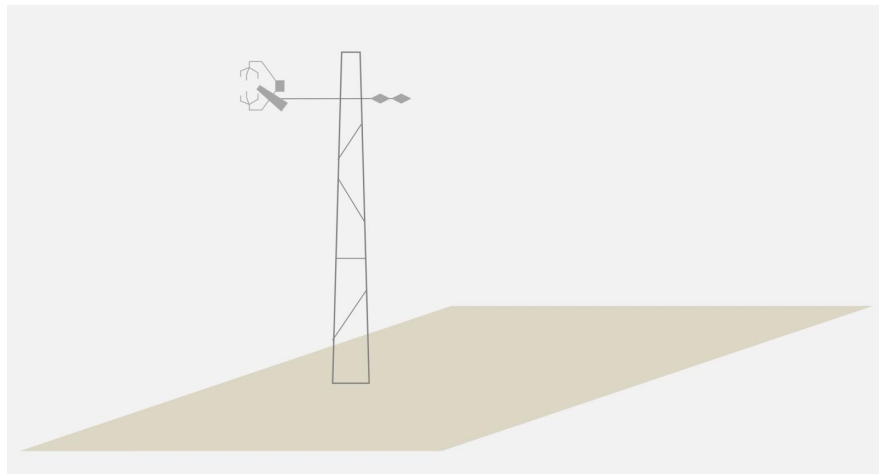
☐ Show Historical Instruments

Customize Columns: 13 Columns Shown

Site Team Variable	CDIAC Description	Early-version BASE Variable	Status	Reset	Delete	FP-In Variable	Reuse values from...	Height (meters)	Instrument Model	Start Date for Height / Instrument Model	Height / Instrument Model Change
FC_WPL_2D	FC_WPL_2D (umol m-2 s-1) CO2 flux with 2D rotation and WPL correction, quality are classified by "Fc_flag"	FC	<input type="checkbox"/>			FC_1_1_1		23.5	GA_OP-LI-COR LI-7500	start of data	Add
			<input type="checkbox"/>						SA-Gill Windmaster Pro		

All sites are different!

AMP needs to know which variables are representative and/or should be aggregated.

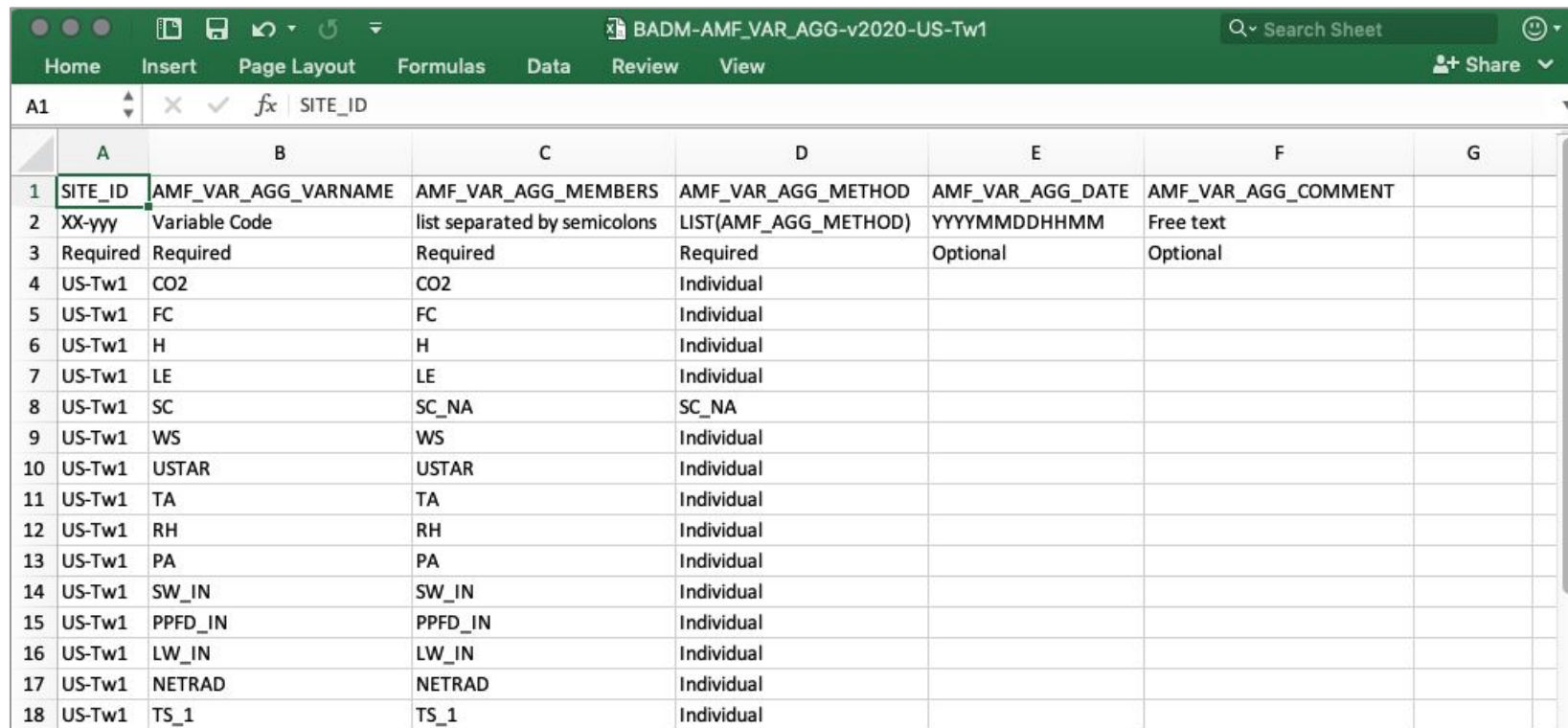


Submitting Variable Aggregation (AMF_VAR_AGG) BADM

1. AMP will contact you requesting submission via email

Submitting Variable Aggregation (AMF_VAR_AGG) BADM

1. AMP will contact you requesting submission via email
2. AMP provides a pre-filled csv file (based on Variable Information and BADM database)



BADM-AMF_VAR_AGG-v2020-US-Tw1						
Home	Insert	Page Layout	Formulas	Data	Review	View
A1	fx	SITE_ID				
	A	B	C	D	E	F
1	SITE_ID	AMF_VAR_AGG_VARNAME	AMF_VAR_AGG_MEMBERS	AMF_VAR_AGG_METHOD	AMF_VAR_AGG_DATE	AMF_VAR_AGG_COMMENT
2	XX-yyy	Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
3	Required	Required	Required	Required	Optional	Optional
4	US-Tw1	CO2	CO2	Individual		
5	US-Tw1	FC	FC	Individual		
6	US-Tw1	H	H	Individual		
7	US-Tw1	LE	LE	Individual		
8	US-Tw1	SC	SC_NA	SC_NA		
9	US-Tw1	WS	WS	Individual		
10	US-Tw1	USTAR	USTAR	Individual		
11	US-Tw1	TA	TA	Individual		
12	US-Tw1	RH	RH	Individual		
13	US-Tw1	PA	PA	Individual		
14	US-Tw1	SW_IN	SW_IN	Individual		
15	US-Tw1	PPFD_IN	PPFD_IN	Individual		
16	US-Tw1	LW_IN	LW_IN	Individual		
17	US-Tw1	NETRAD	NETRAD	Individual		
18	US-Tw1	TS_1	TS_1	Individual		

Submitting Variable Aggregation (AMF_VAR_AGG) BADM

1. AMP will contact you requesting submission via email
2. AMP provides a pre-filled csv file (based on Variable Information and BADM database)
3. AMP provides instructions in the email

AmeriFlux TESTQAQC-3339 ACTION REQUIRED: Submit ONEFlux variable aggregation info

Danielle Christianson amp.data.qaac@gmail.com via berkeley.edu
to dsvehla

Thu, Sep 3, 1:21 PM

Dear Gilberto, You-Wei,

We are preparing to process US-ARM data using the [ONEFlux processing pipeline](#), the eddy covariance data processing codes package used create uniform gap-filled products, uncertainty estimates, and partitioned CO2 fluxes data products for AmeriFlux, other regional flux networks and for FLUXNET.

To process the site's data, we need more information on which variables submitted to AmeriFlux are representative for the site and/or how we should aggregate these variables for processing.

To process the site's data, we need more information on the variables that are representative for the site, and also information on aggregating these variables for processing.

The attached file with proposed variable aggregation information is generated from our database.

Please confirm and correct any errors by completing the following:

1. Review the attached proposed variable aggregation information (CSV file).
2. Correct any errors. See QuickView instructions and additional explanations below.
3. Upload the CSV file at the [Upload Data](#) page using the BADM tab (login required).

QuickView instructions:

Varname

Variable to be used in ONEFlux processing.

MEMBERS

MEMBERS variable(s) must match variables in the site's BASE data product.

Indicate storage with VARNAME = SC

Options for MEMBERS:

- Enter SC variable(s) calculated by site team
- Enter CO2 variable to use for estimating SC (typically the highest CO2 observation)
- Storage is negligible. Enter SC_NA.

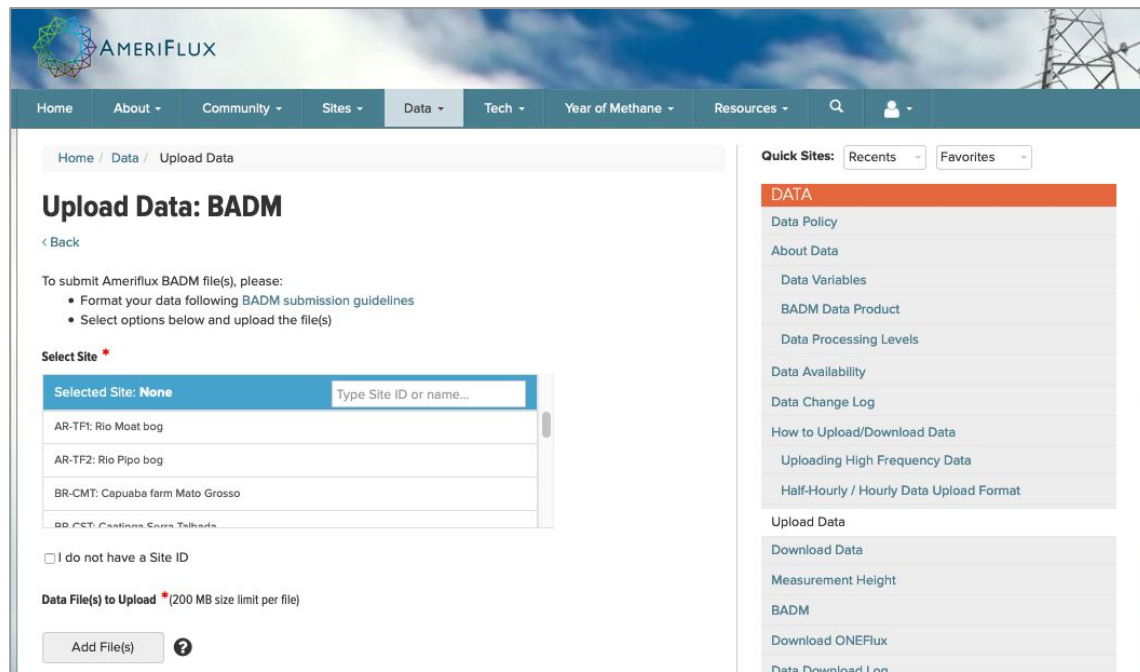
AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
CO2	CO2_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		

DATE

- No DATE = info starts at beginning of data record.
- Provide DATE when info changes within data record.

Submitting Variable Aggregation (AMF_VAR_AGG) BADM

1. AMP will contact you requesting submission via email
2. AMP provides a pre-filled csv file (based on Variable Information and BADM database)
3. AMP provides instructions in the email
4. Site Teams review, update if needed, and upload the csv file



The screenshot shows the AmeriFlux website interface. The top navigation bar includes links for Home, About, Community, Sites, Data, Tech, Year of Methane, and Resources. The main content area is titled 'Upload Data: BADM' and includes a '< Back' link. Below the title, instructions state: 'To submit Ameriflux BADM file(s), please:'. A list of instructions follows: 'Format your data following BADM submission guidelines' and 'Select options below and upload the file(s)'. A 'Select Site' section features a dropdown menu with 'Selected Site: None' and a search box 'Type Site ID or name...'. The dropdown list shows several sites: 'AR-TF1: Rio Moat bog', 'AR-TF2: Rio Pipo bog', 'BR-CMT: Capuba farm Mato Grosso', and 'BR-CST: Canteiro Serra Talhada'. Below the dropdown, there is a checkbox labeled 'I do not have a Site ID'. At the bottom, there is a section 'Data File(s) to Upload' with a note '(200 MB size limit per file)' and an 'Add File(s)' button with a help icon. On the right side, there is a 'Quick Sites' section with 'Recents' and 'Favorites' tabs, and a 'DATA' section with a list of links: 'Data Policy', 'About Data', 'Data Variables', 'BADM Data Product', 'Data Processing Levels', 'Data Availability', 'Data Change Log', 'How to Upload/Download Data', 'Uploading High Frequency Data', 'Half-Hourly / Hourly Data Upload Format', 'Upload Data', 'Download Data', 'Measurement Height', 'BADM', 'Download ONEFlux', and 'Data Download Log'.

<https://ameriflux.lbl.gov/data/upload-data/> > Choose BADM

Review Variable Aggregation (AMF_VAR_AGG) csv file

AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
C02	C02_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		
G_1	G_1_1_1;G_2_1_1;G_3_1_1;G_4_1_1	Mean		
SWC_1	SWC_1_1_1;SWC_2_1_1;SWC_2_1_2	Mean		AMP: No depth info; please double check

Review Variable Aggregation (AMF_VAR_AGG) csv file

Varname

Variable to be used in ONEflux processing.

MEMBERS

MEMBERS variable(s) must match variables in the site's BASE data product.

AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
C02	C02_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		
G_1	G_1_1_1;G_2_1_1;G_3_1_1;G_4_1_1	Mean		
SWC_1	SWC_1_1_1;SWC_2_1_1;SWC_2_1_2	Mean		AMP: No depth info; please double check

METHOD Options

- Individual
- Mean
- SC_NA (use only with SC_NA in MEMBERS)

Review Variable Aggregation (AMF_VAR_AGG) csv file

Varname

Variable to be used in ONEflux processing.

MEMBERS

MEMBERS variable(s) must match variables in the site's BASE data product.

AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
C02	C02_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		
G_1	G_1_1_1;G_2_1_1;G_3_1_1;G_4_1_1	Mean		
SWC_1	SWC_1_1_1;SWC_2_1_1;SWC_2_1_2	Mean		AMP: No depth info; please double check

METHOD Options

- Individual
- Mean
- SC_NA (use only with SC_NA in MEMBERS)

Review Variable Aggregation (AMF_VAR_AGG) csv file

Varname

Variable to be used in ONEflux processing.

MEMBERS

MEMBERS variable(s) must match variables in the site's BASE data product.

AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
C02	C02_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		
G_1	G_1_1_1;G_2_1_1;G_3_1_1;G_4_1_1	Mean		
SWC_1	SWC_1_1_1;SWC_2_1_1;SWC_2_1_2	Mean		AMP: No depth info; please double check

DATE

- No DATE = info starts at beginning of data record.
- Provide DATE when info changes within data record.

METHOD Options

- Individual
- Mean
- SC_NA (use only with SC_NA in MEMBERS)

Review Variable Aggregation (AMF_VAR_AGG) csv file

Varname

Variable to be used in ONEFlux processing.

MEMBERS

MEMBERS variable(s) must match variables in the site's BASE data product.

Indicate storage with VARNAME = SC

Options for MEMBERS:

- Enter SC variable(s) calculated by site team
- Enter C02 variable to use for estimating SC (typically the highest C02 observation)
- Storage is negligible. Enter SC_NA.

See details below for more info.

AMF_..._VARNAME	AMF_..._MEMBERS	AMF_..._METHOD	AMF_..._DATE	AMF_..._COMMENT
Variable Code	list separated by semicolons	LIST(AMF_AGG_METHOD)	YYYYMMDDHHMM	Free text
Required	Required	Required	Optional	Optional
C02	C02_1_1_1	Individual		
FC	FC_1_1_1	Individual		
SC	SC_NA	SC_NA		
TA	TA_1_1_1	Individual		
TA	TA_1_1_1;TA_2_1_1	Mean	201504121200	
WS	WS_1_1_1	Individual		
G_1	G_1_1_1;G_2_1_1;G_3_1_1;G_4_1_1	Mean		
SWC_1	SWC_1_1_1;SWC_2_1_1;SWC_2_1_2	Mean		AMP: No depth info; please double check

DATE

- No DATE = info starts at beginning of data record.
- Provide DATE when info changes within data record.

METHOD Options

- Individual
- Mean
- SC_NA (use only with SC_NA in MEMBERS)

Reporting carbon storage (VARNAME = SC)

Storage is important

OR

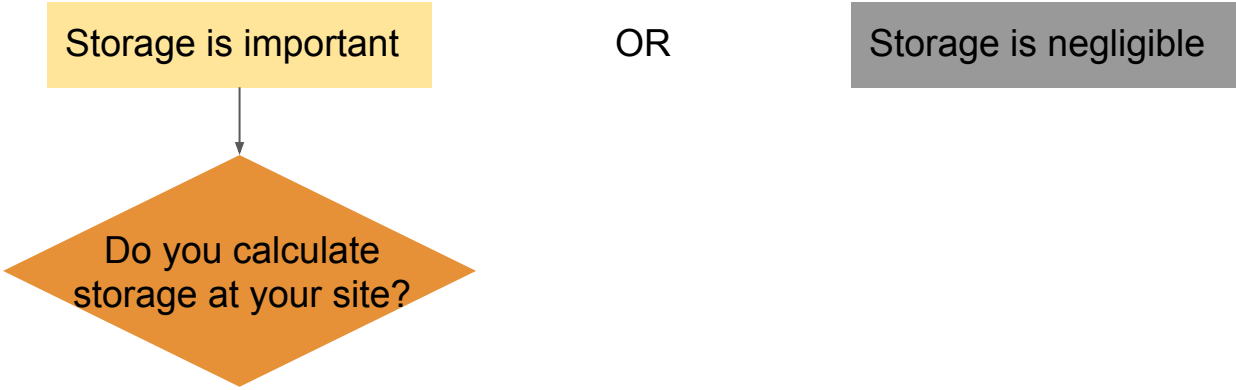
Storage is negligible

Reporting carbon storage (VARNAME = SC)

Storage is important

OR

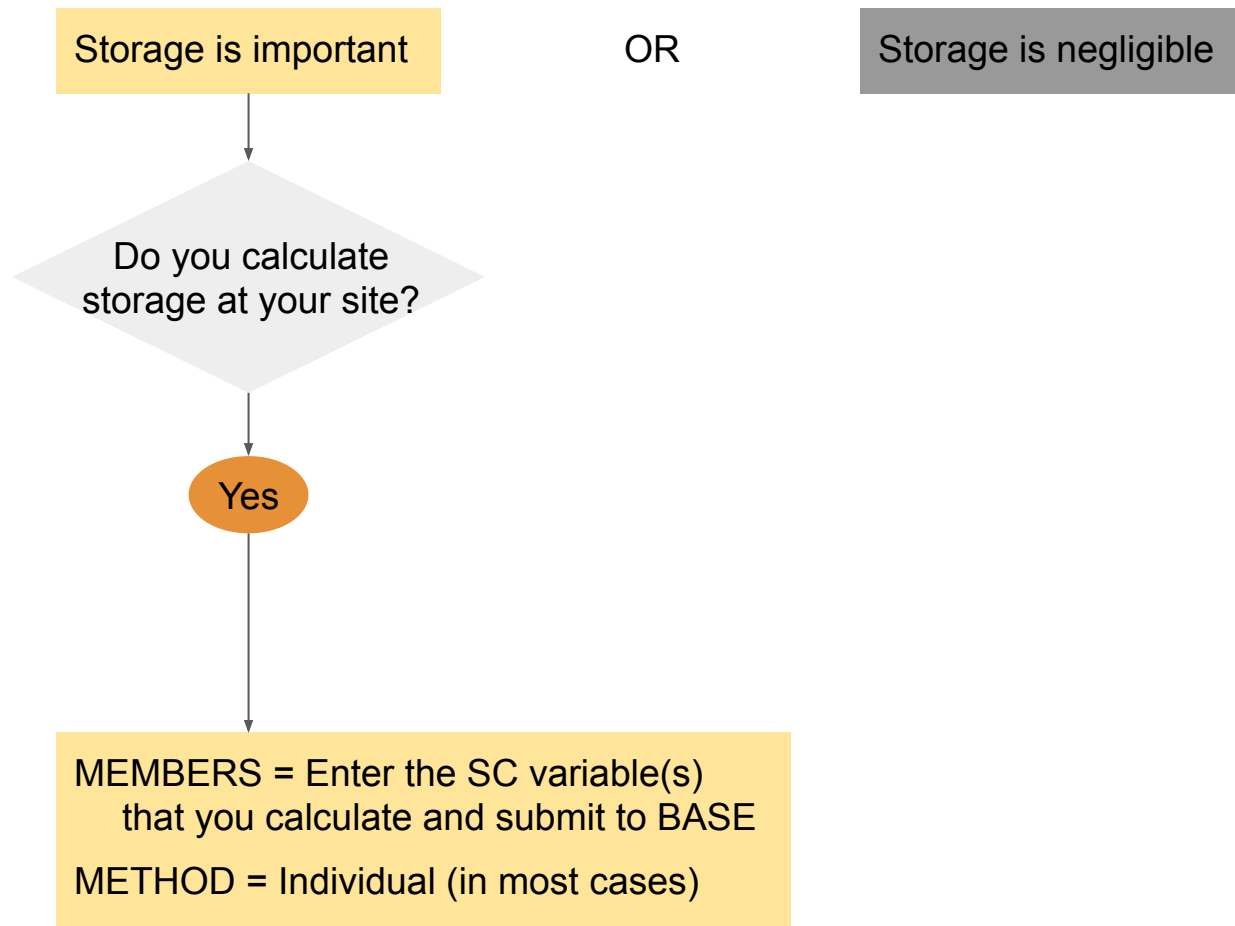
Storage is negligible



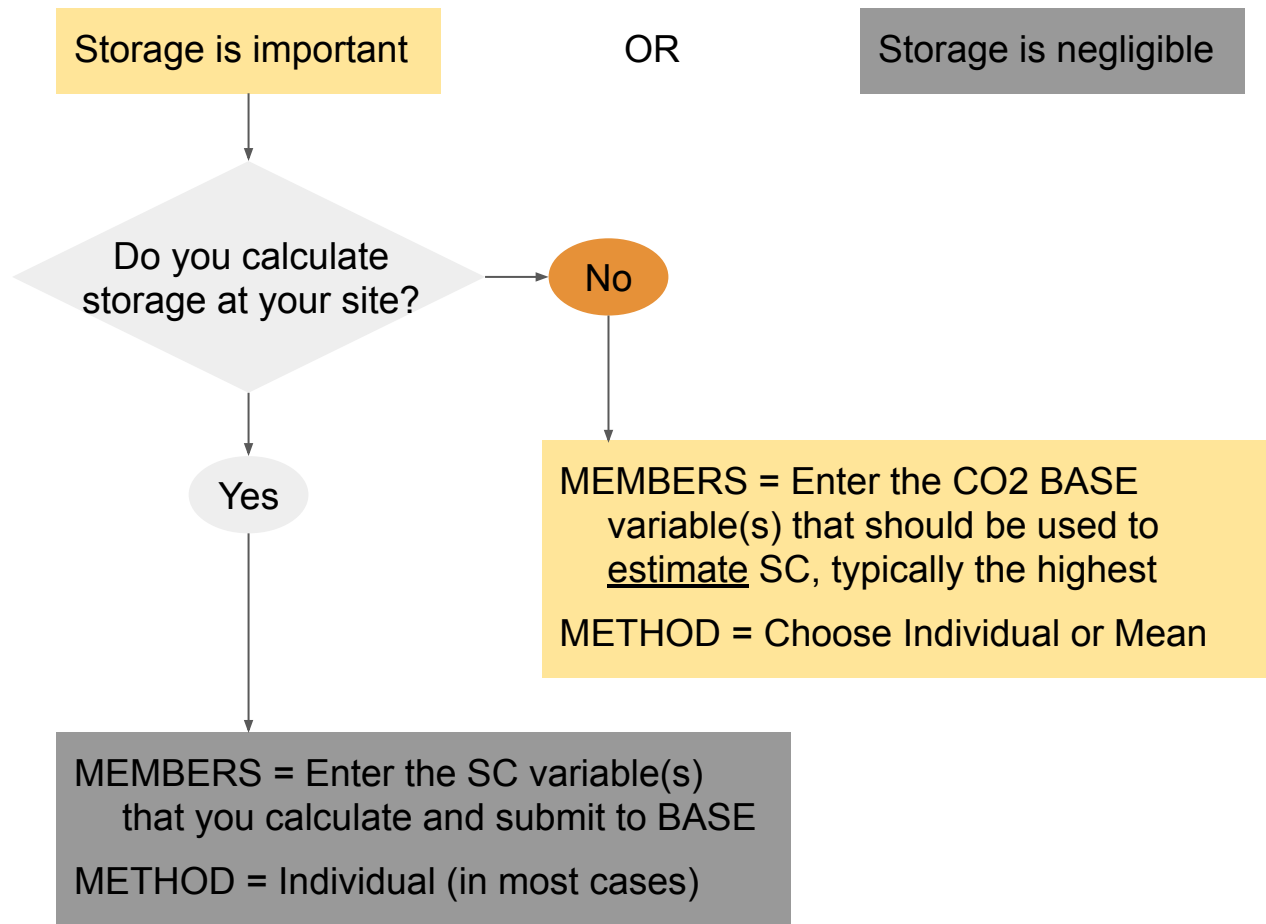
```
graph TD; A[Storage is important] --> D{Do you calculate storage at your site?}; B[Storage is negligible];
```

Do you calculate
storage at your site?

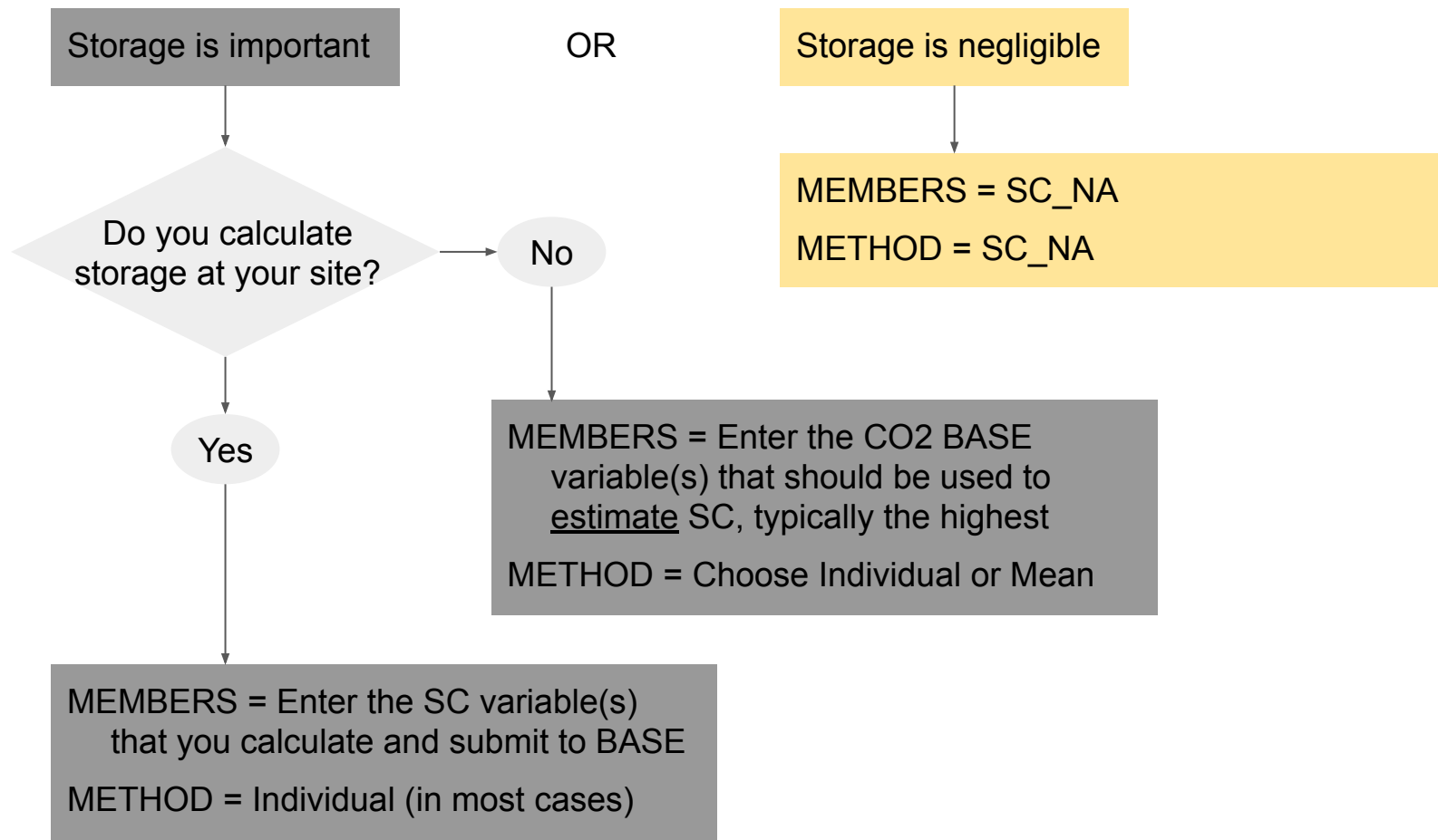
Reporting carbon storage (VARNAME = SC)



Reporting carbon storage (VARNAME = SC)



Reporting carbon storage (VARNAME = SC)



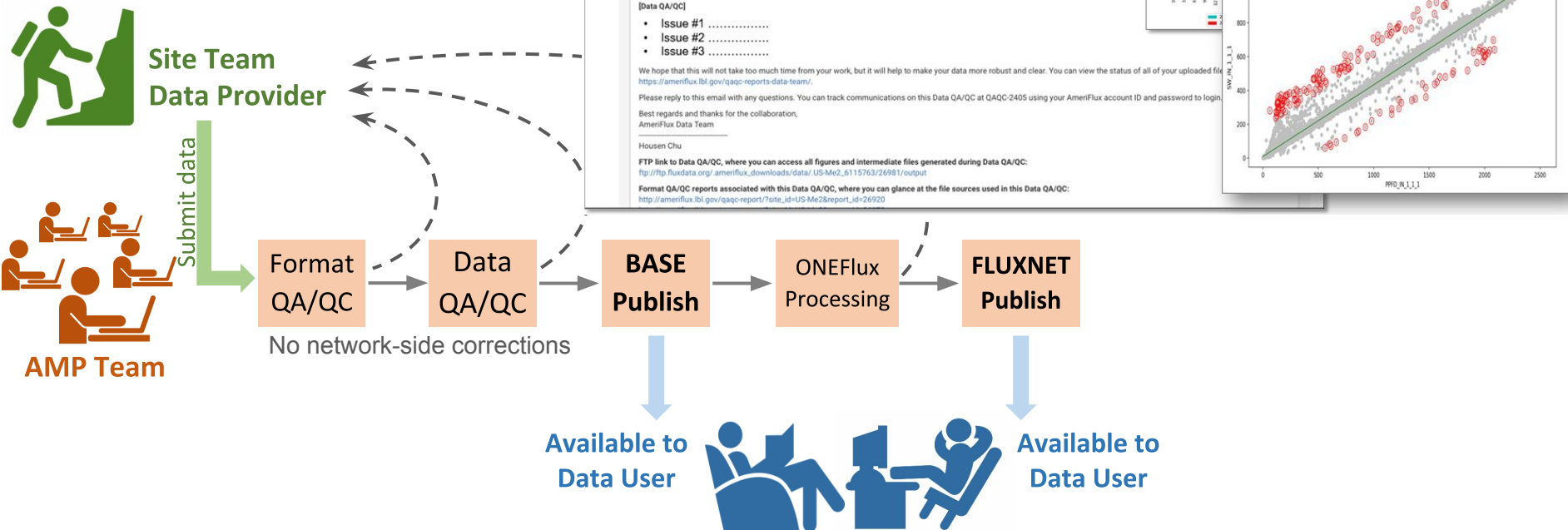
Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products



Housen
Chu

AmeriFlux Data Pipeline





Observational data patterns for time series data quality assessment

Pastorello, G. et al., 2014., Proc. 10th IEEE International Conference on e-Science, Sao Paulo, pp. 271-278, doi: 10.1109/eScience.2014.45

Issues

Descriptions

Sources

Corrections

Windowed varying maximums

levels of values at higher end varies more than expected for natural variability

sensor malfunction or dirt accumulation; poor filtering for noise and spikes

correct calibration; redo de-spiking from original data

Trend in minimums

noticeable increasing or decreasing trend in lower end of vales within a time window

sensor degradation

Trend in maximums

noticeable increasing or decreasing trend in higher end of vales within a time window

sensor degradation

Measurement interval filter

data filtered for one or more ranges of values (e.g., values between x and y are removed)

incorrect filtering process

Measurement cutoff filter

data at points higher (or lower) than a threshold are removed (e.g., values larger than x are removed)

incorrect filtering process

Inconsistent noise levels

time series has different levels of noise that are not usually explained by sensor, measurement, or processing variability, especially very low noise periods

indirect estimation based gap-processing filtering

Repeating pattern

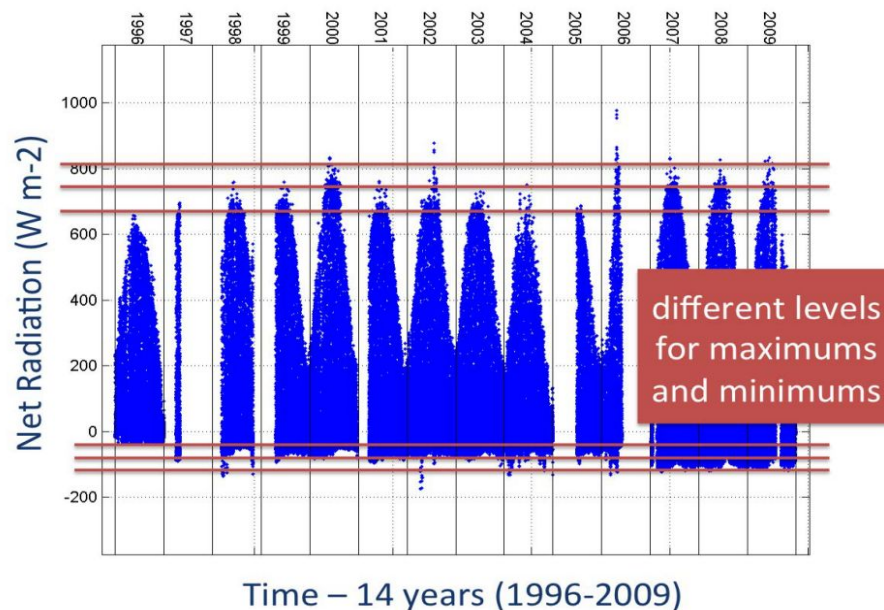
repetition of exact pattern throughout short or long time windows

indirect estimation based gap-processing

Physically unlikely ranges

range of values for a period within available data has unlikely values (either because it's out of physically possible values or because it grossly mismatches general trend)

sensor malfunction; inefficient gap-processing filtering



Current Implemented Checks

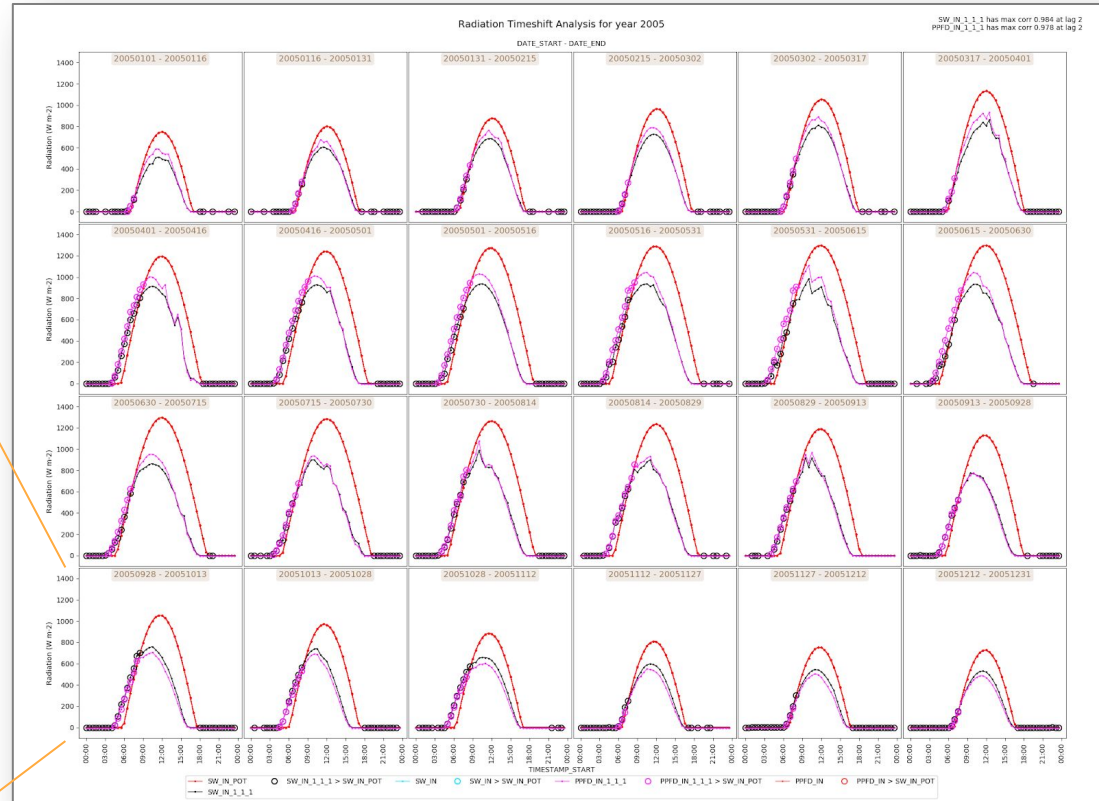
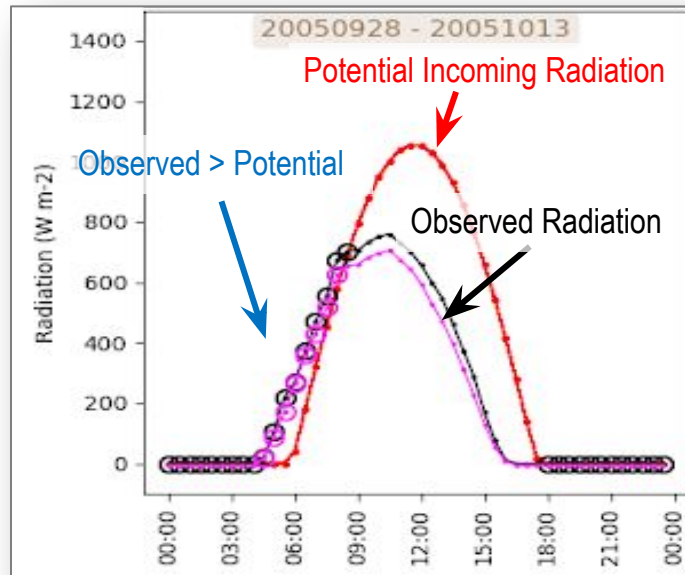
- Timestamp shift check
- Physical range check
- Diurnal & Seasonal pattern
- Multivariate comparison
- USTAR filtering check
- Variable availability

Current Implemented Checks

o Timestamp shift

- Timestamp shift
- Daylight saving time
- Data stream not synchronized
- Time zone specification

Max Diurnal Pattern

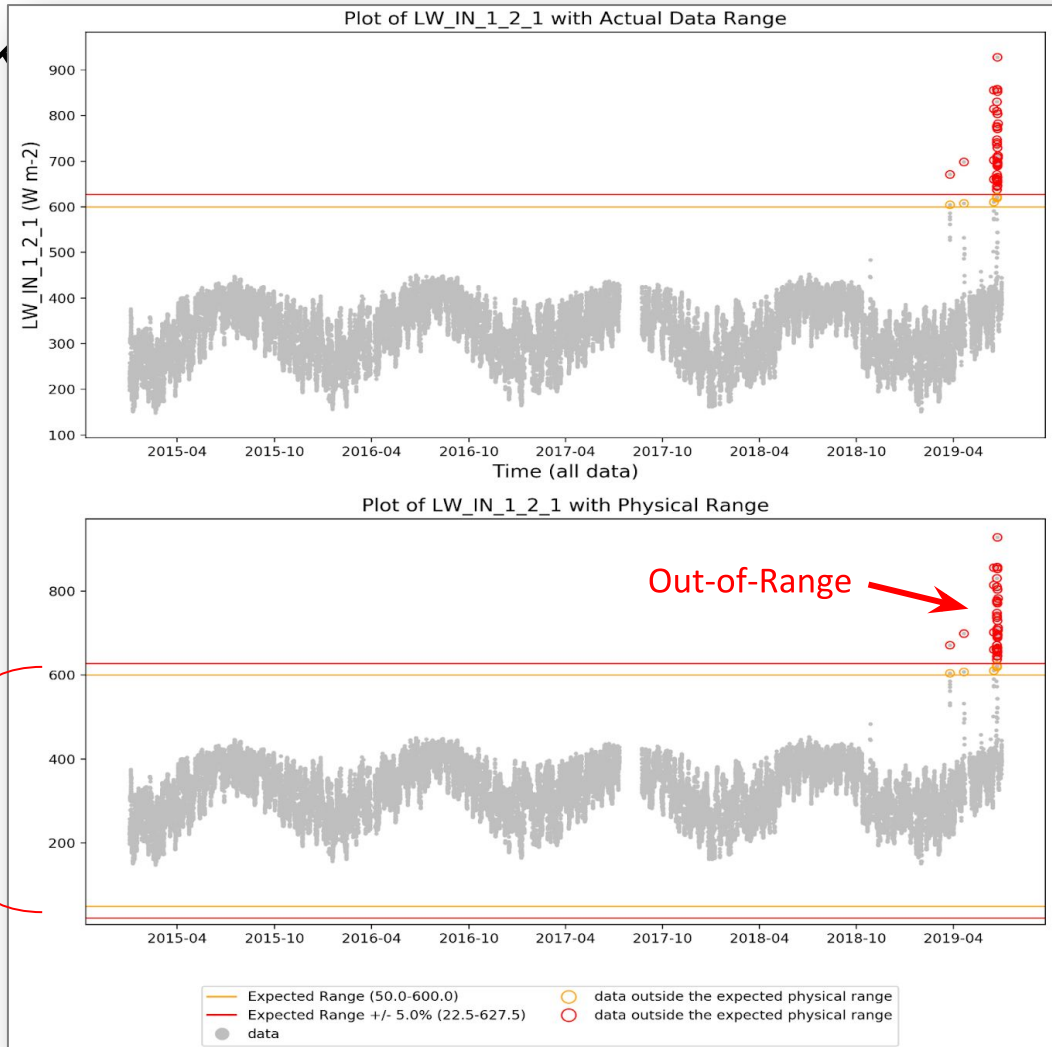


Current Implemented Check

o Physical range check

- Physically unlikely values
- Outlier
- Inconsistent noise levels / filtering
- Trend
- Step change in full range
- Change of resolution
- Repeating patterns / constants
- Long gaps
- Measurement cutoff filter
- Other unrecognized pattern

Physically plausible range
Network-wise historical range

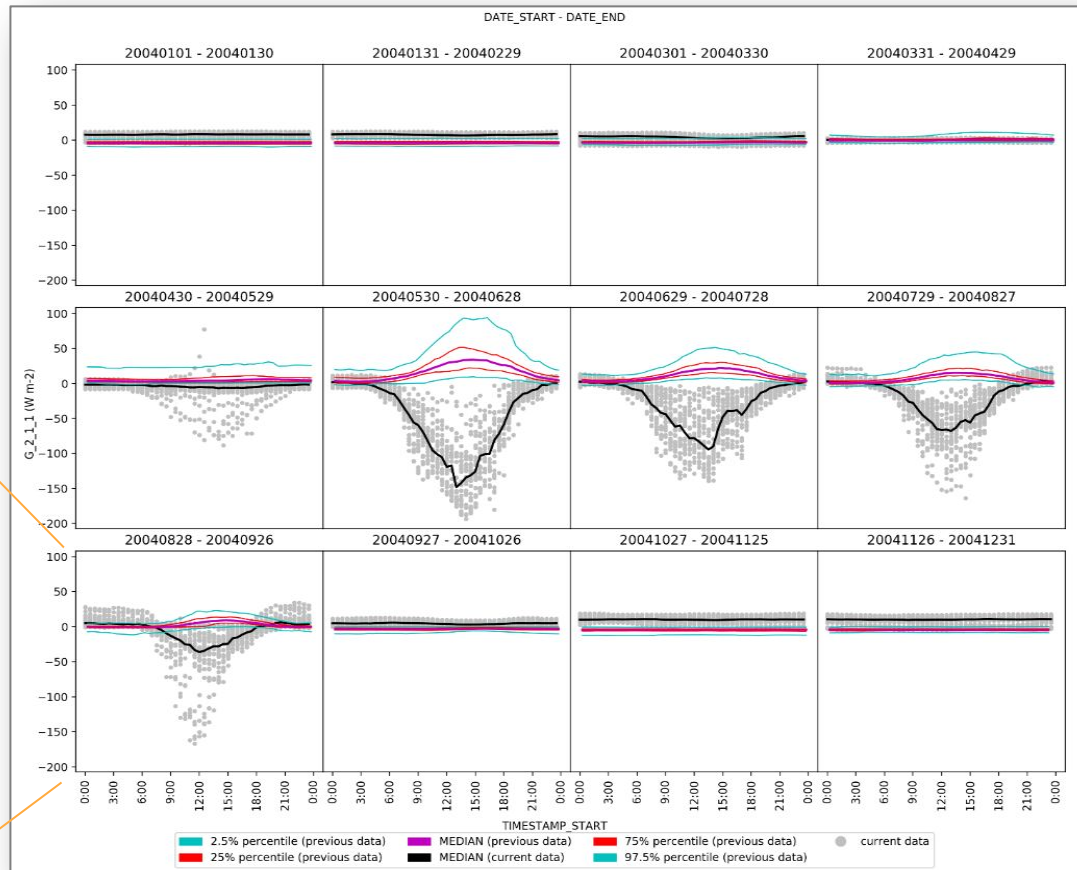
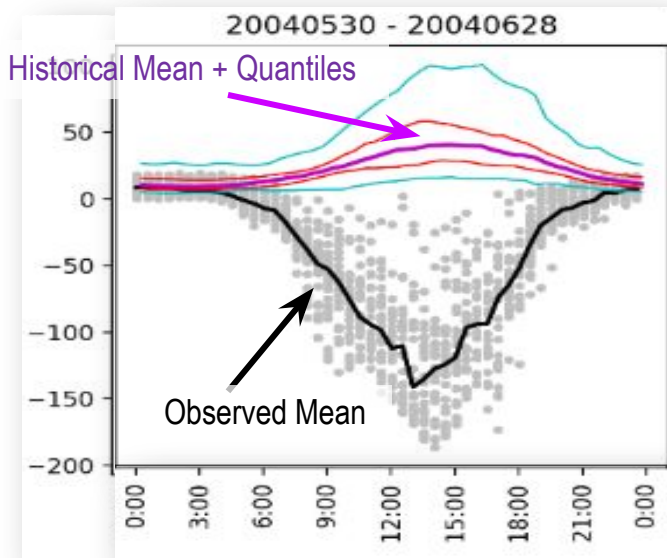


Current Implemented Checks

o Diurnal & Seasonal pattern

- Physically unlikely values
- Outlier
- Sign convention
- Step change in full range

Mean Diurnal Pattern



Current Implemented Checks

o Multivariate comparison

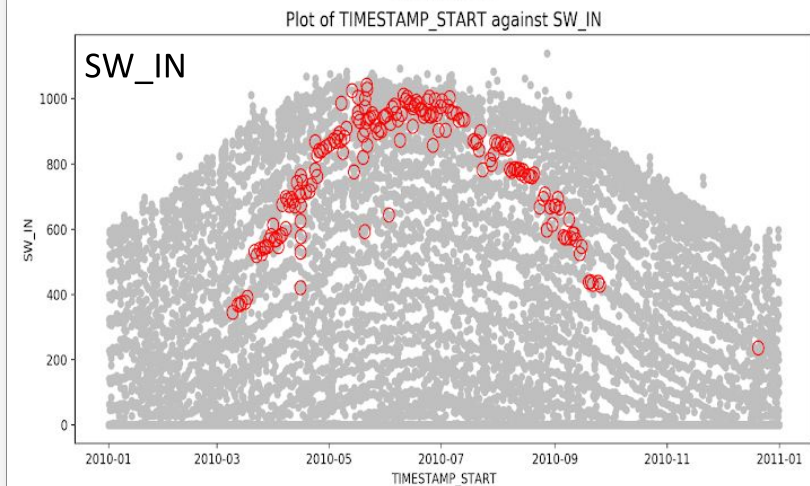
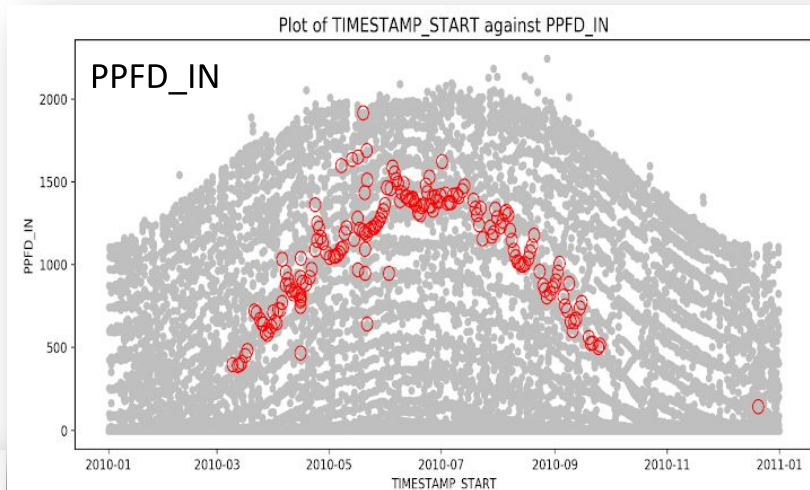
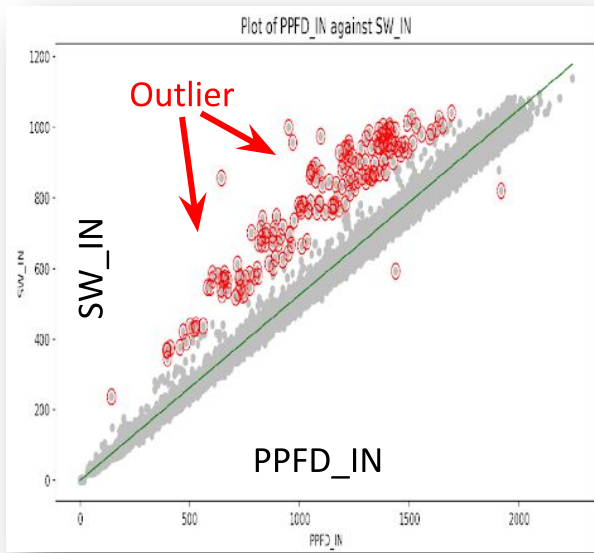
- Outlier
- Variables not synchronized in time
- Step change in full range
- Trend
- Shaded radiation
- Derived one from other

• Cross-variable:

- PPFD_IN vs SW_IN
- TA vs T_SONIC
- WS vs USTAR

• Cross-level:

- TA profile

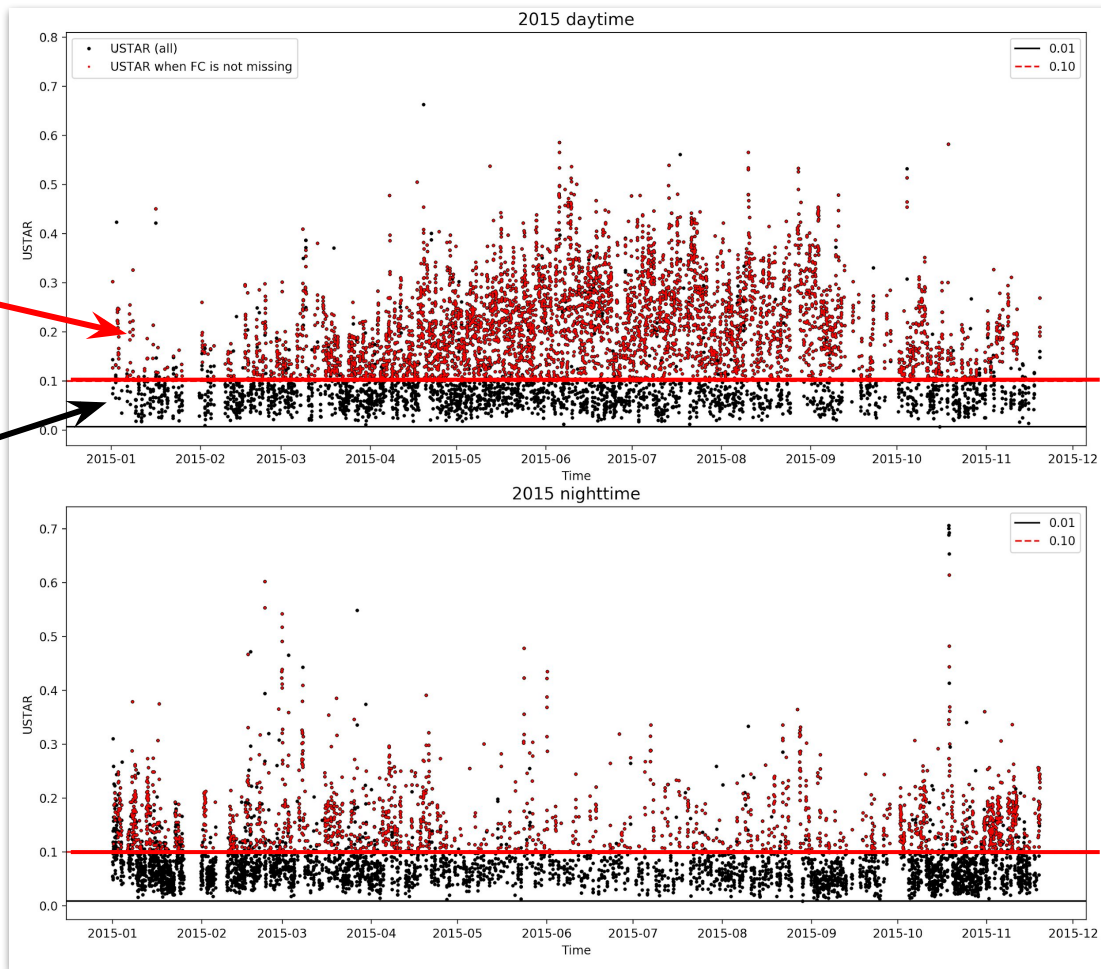


Current Implemented Checks

- FC-USTAR filtering check

USTAR when FC not
missing

USTAR when FC
missing

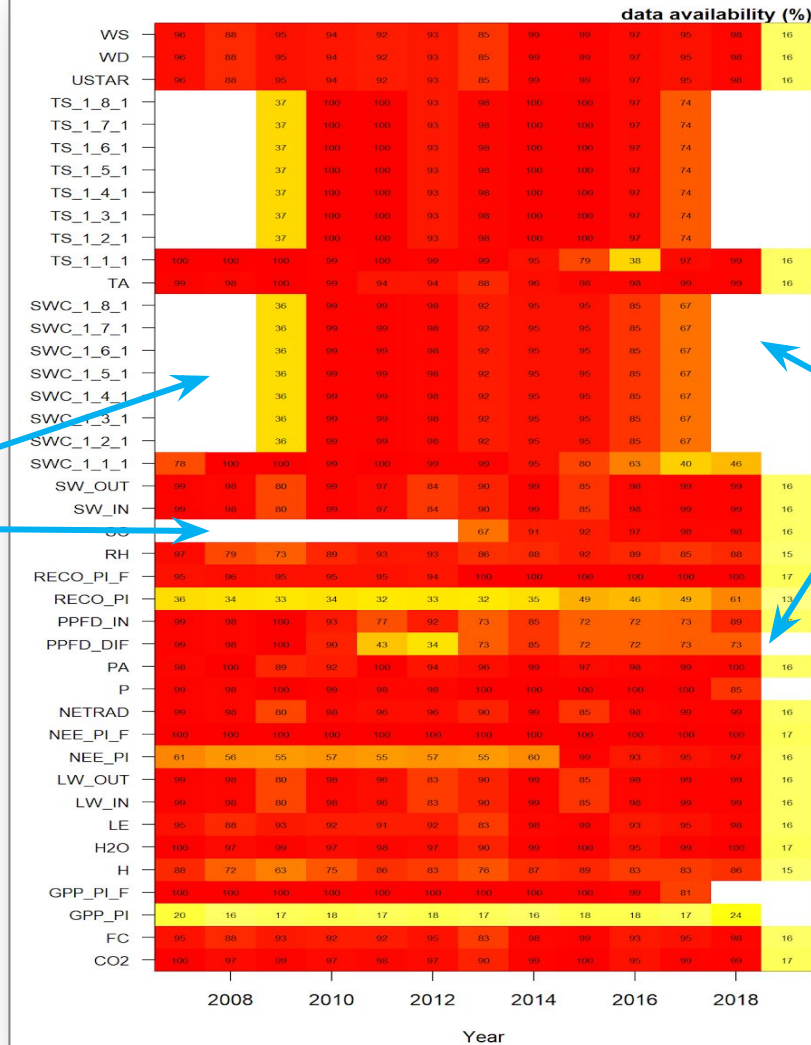


Current Implemented Checks

- o Variable availability
 - Long gaps
 - Missing mandatory variables
 - Inconsistent variable naming / qualifier / aggregation

No measurement

Variables not provided



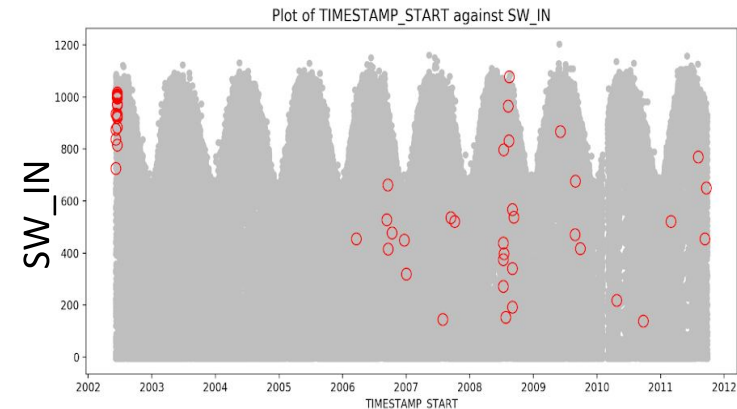
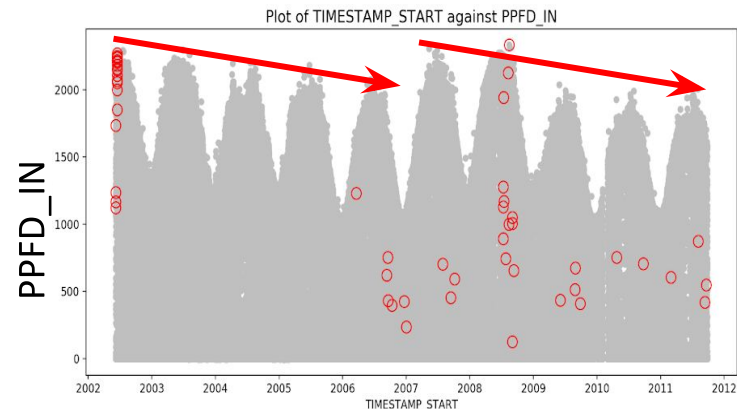
Data QA/QC for ONEFlux

- Important variables

Required	Recommended	Suggested
FC	NETRAD	WD
CO2	G	P
SC		PPFD_IN
USTAR		SW_OUT
WS		LW_IN
SW_IN		LW_OUT
TA		SWC ...
RH		TS ...
LE		
H		
PA		

- Critical issues affecting processing

- Timestamp shifts
- Flux variables being filtered using USTAR thresholds
- Trend or step change in radiation variables
- Large outliers, esp in flux variables



Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- **FLUXNET data products created with ONEFlux**
- Future of FLUXNET regional and global data products



Gilberto
Pastorello

FLUXNET data products created with ONEFlux

H

LE

NEE

RECO

GPP

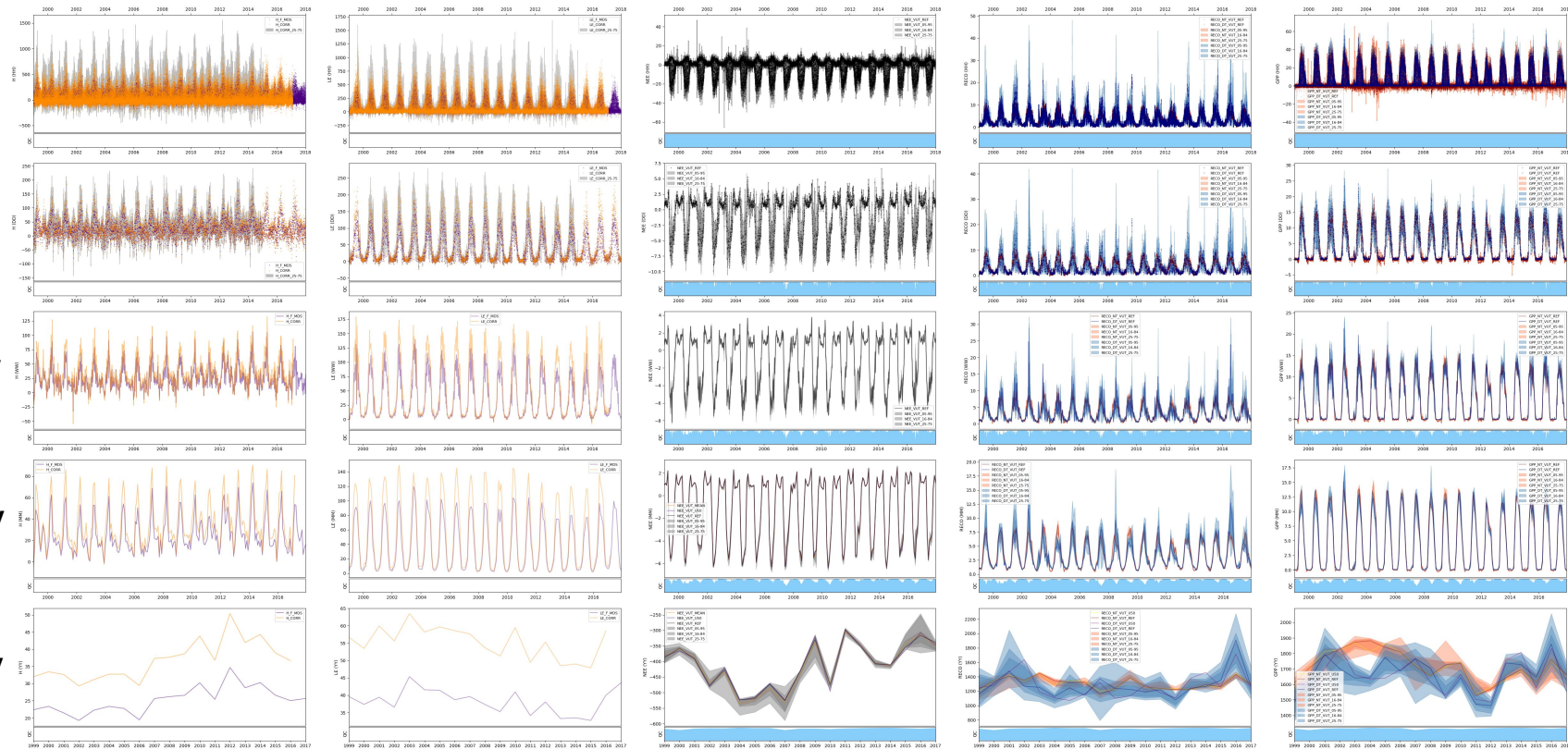
(half-) hourly

daily

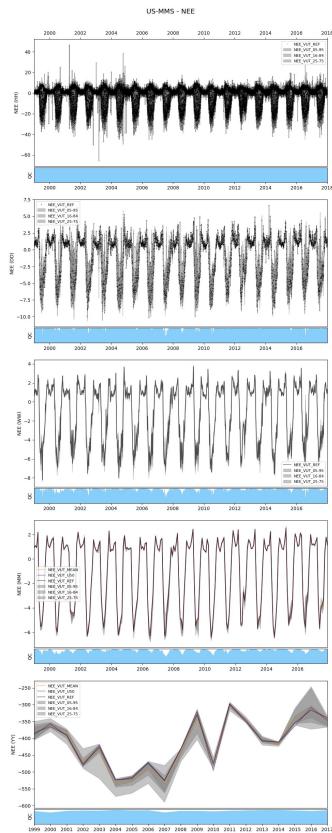
weekly

monthly

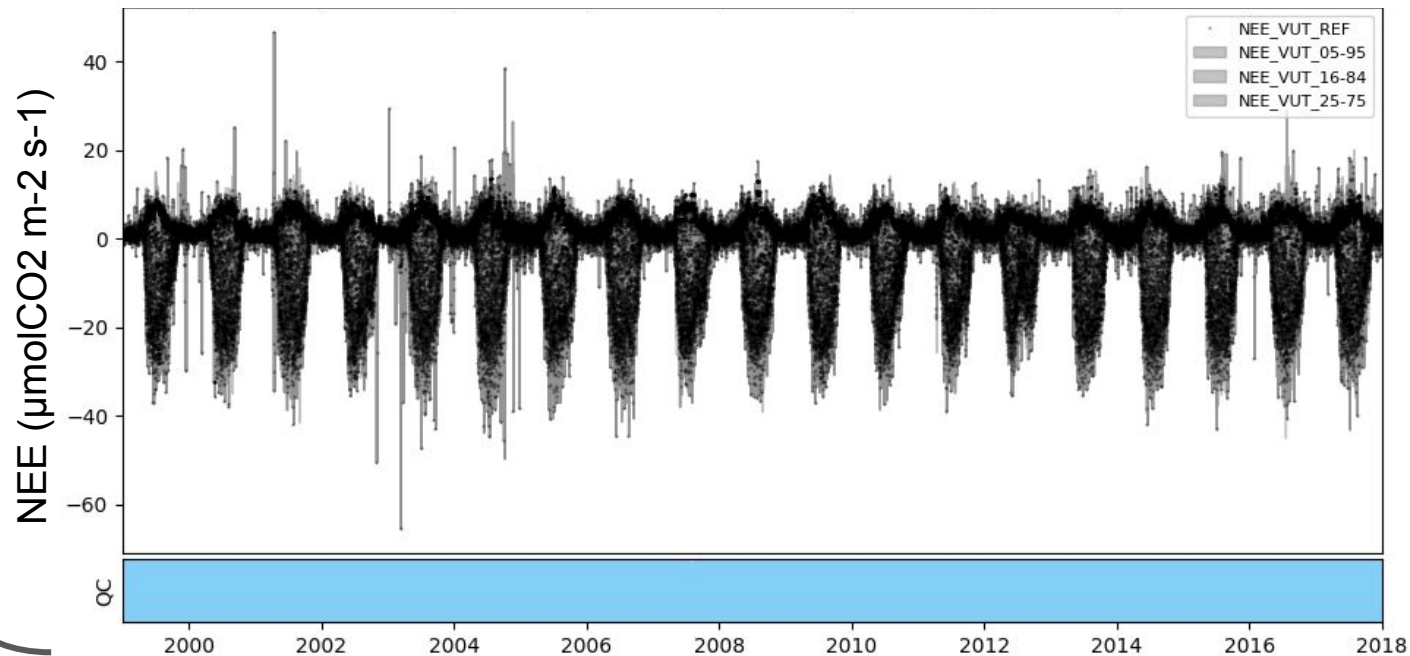
yearly



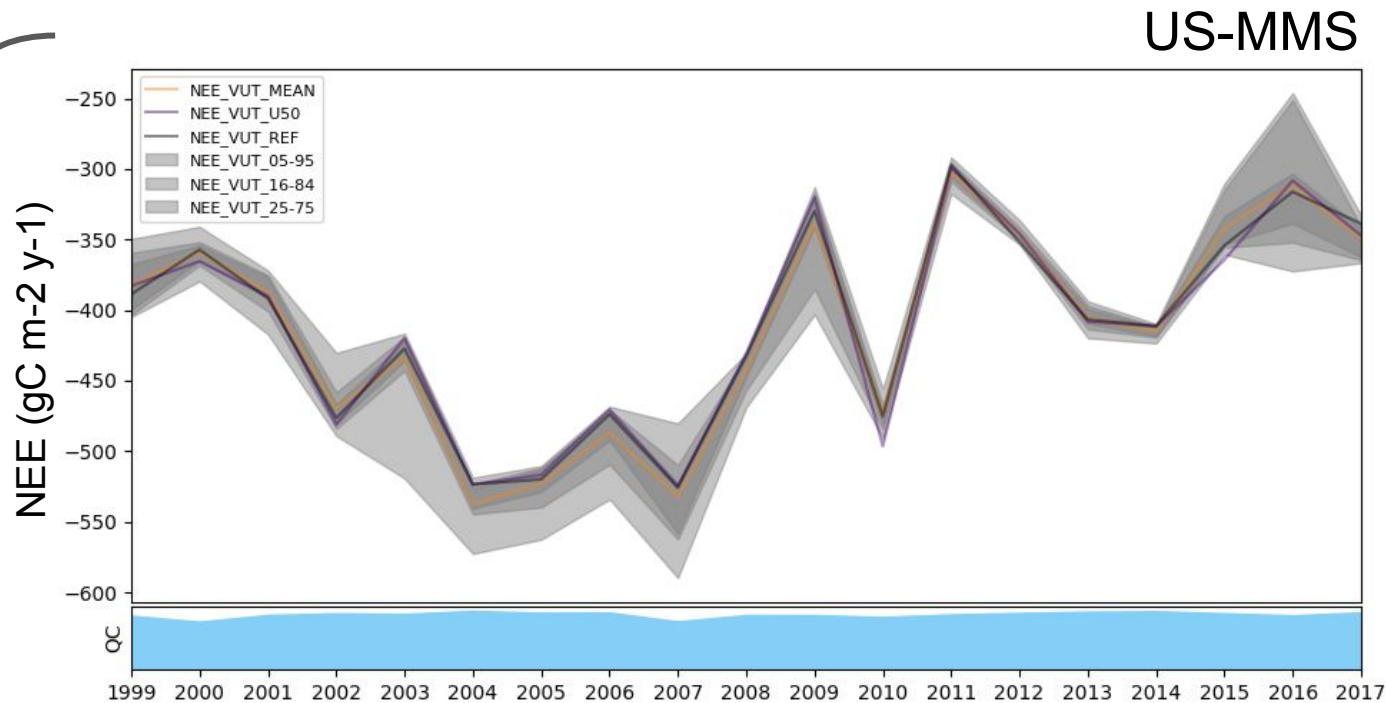
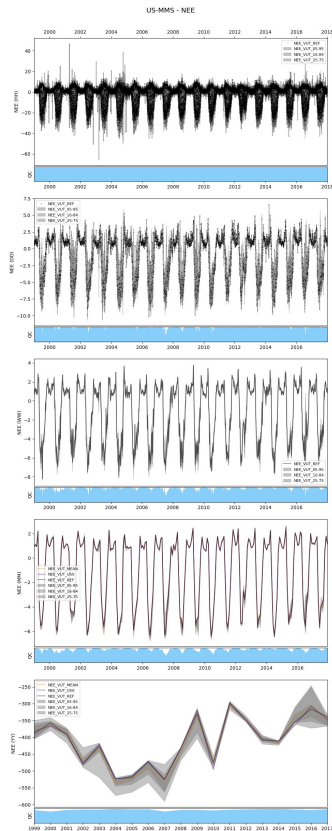
NEE ($\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$), hourly



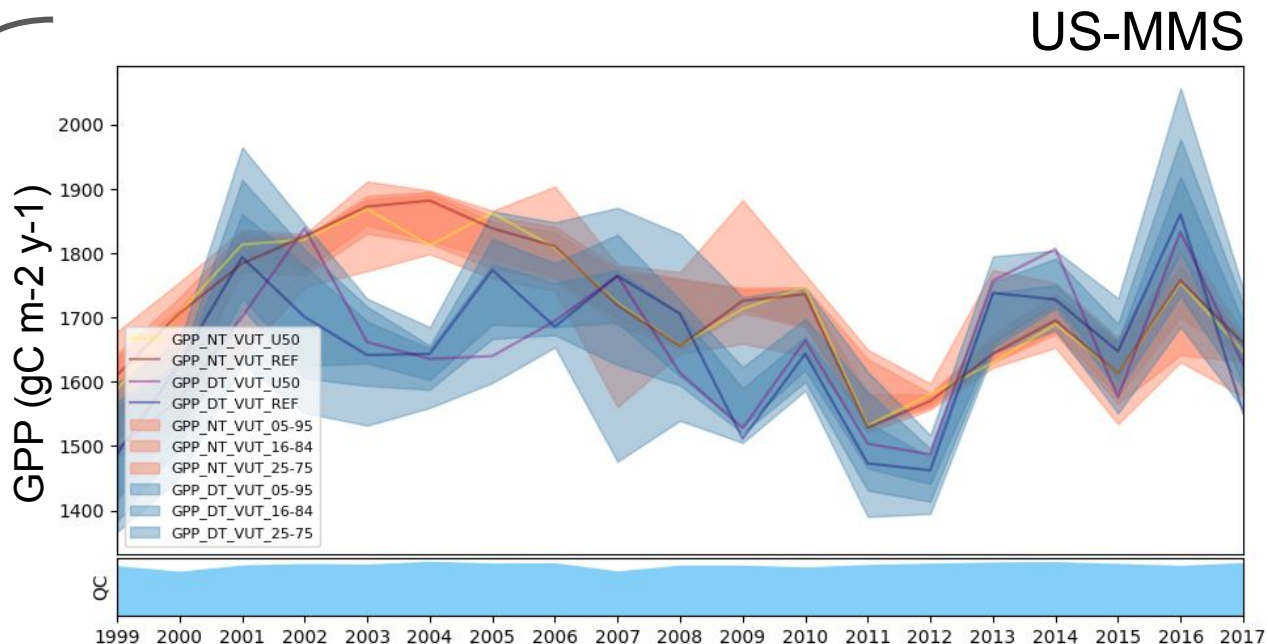
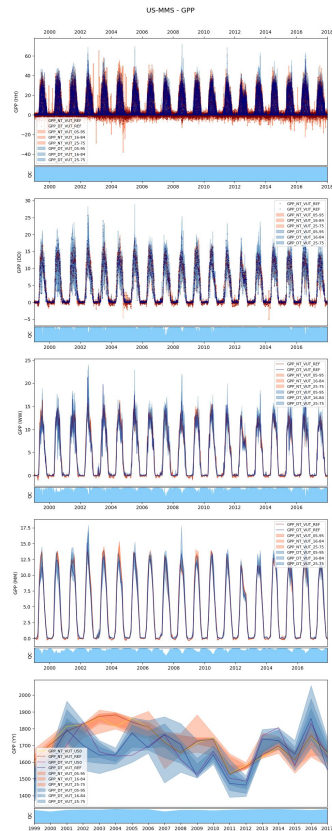
US-MMS



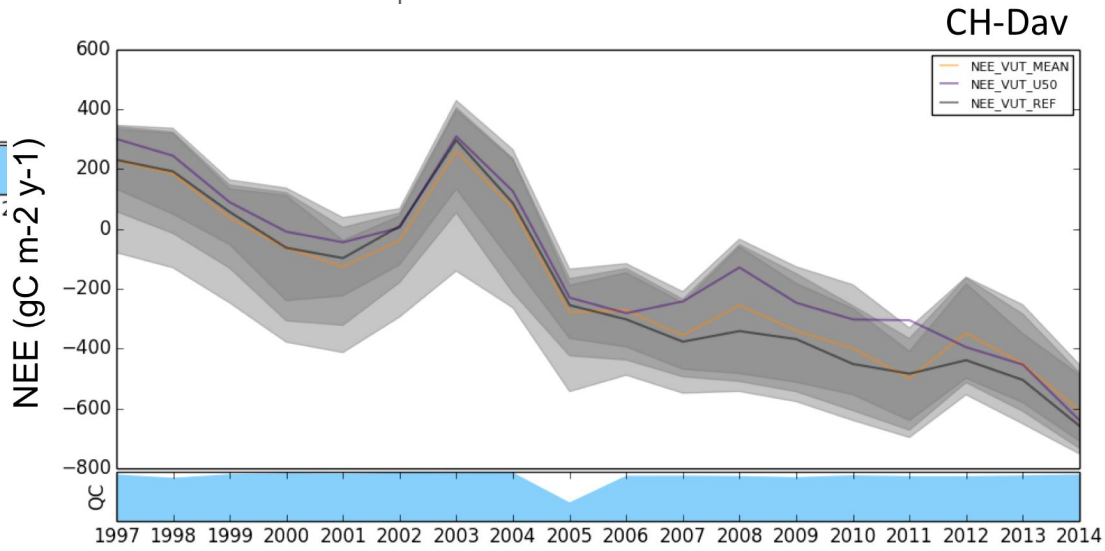
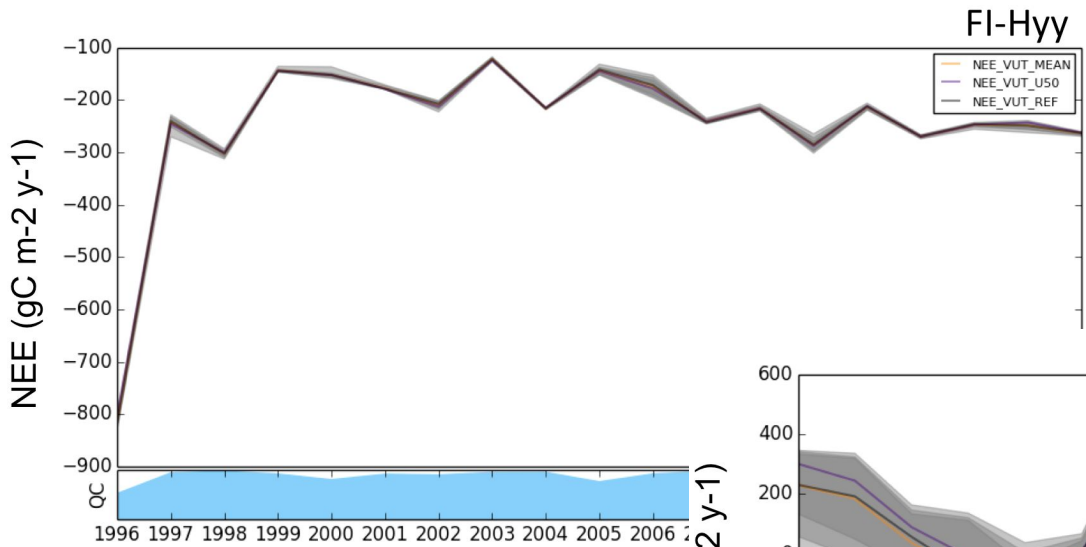
NEE (gC m⁻² y⁻¹), yearly



GPP (gC m⁻² y⁻¹), yearly

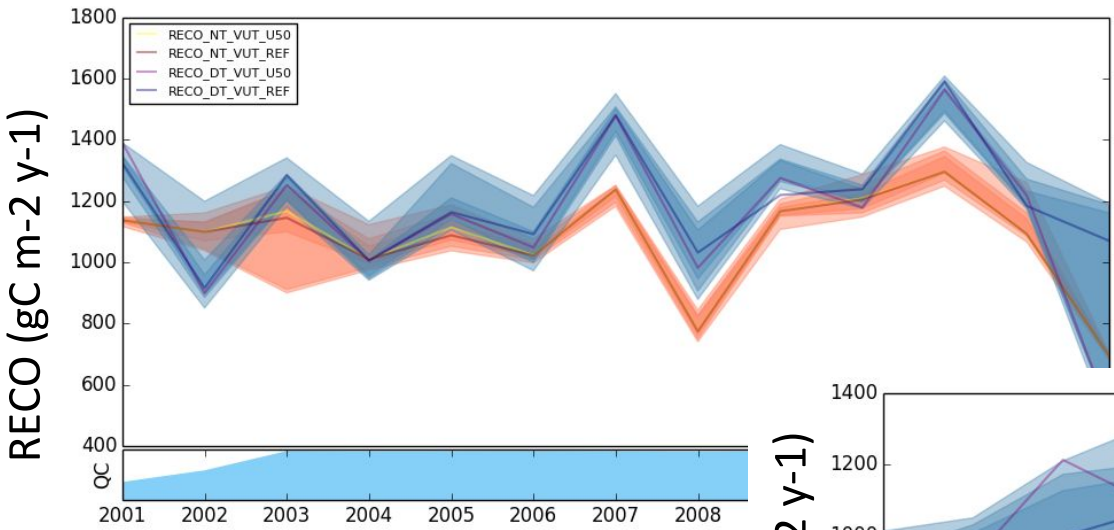


NEE (gC m-2 y-1), yearly

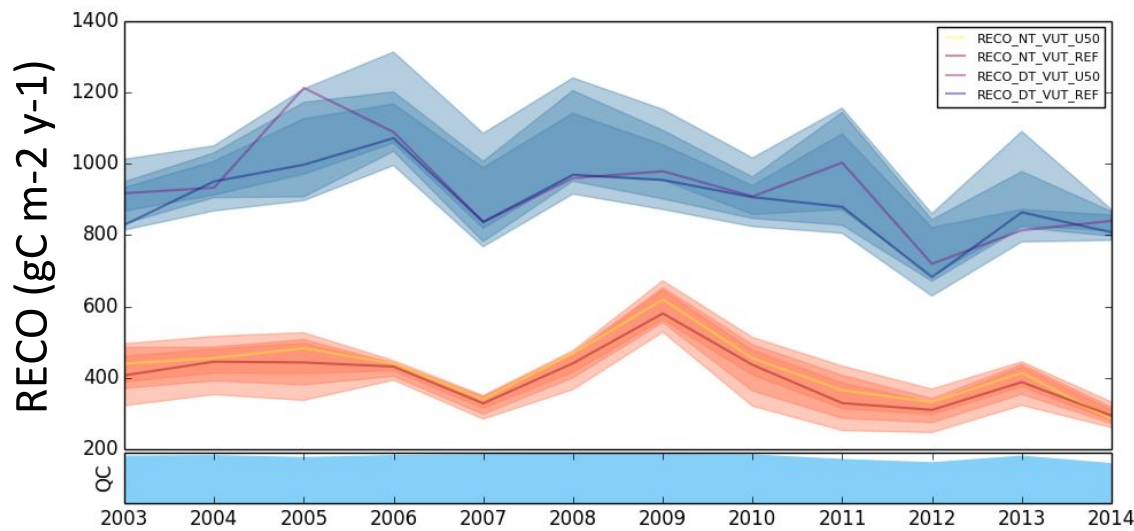


RECO (gC m-2 y-1), yearly

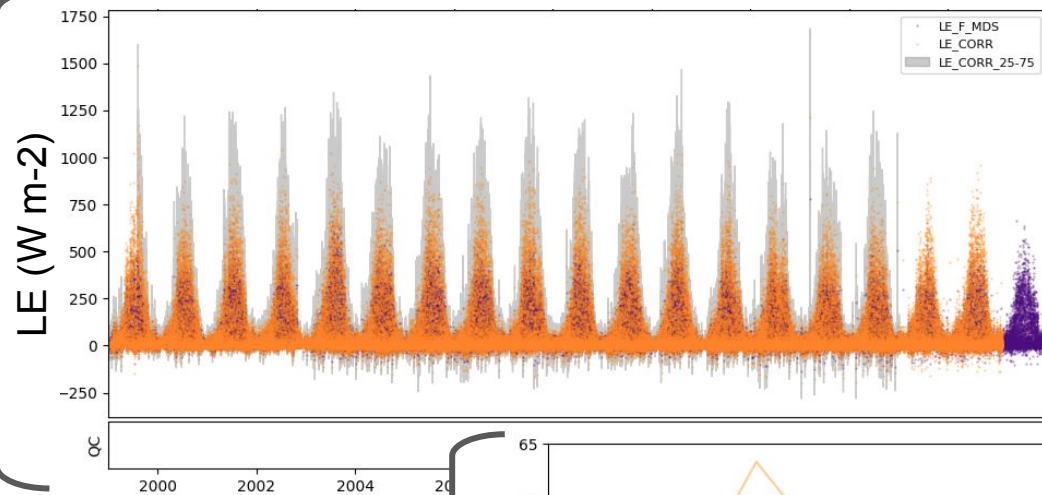
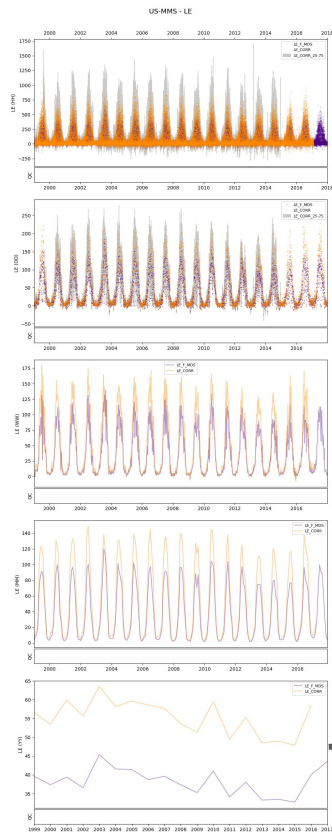
US-Ne2



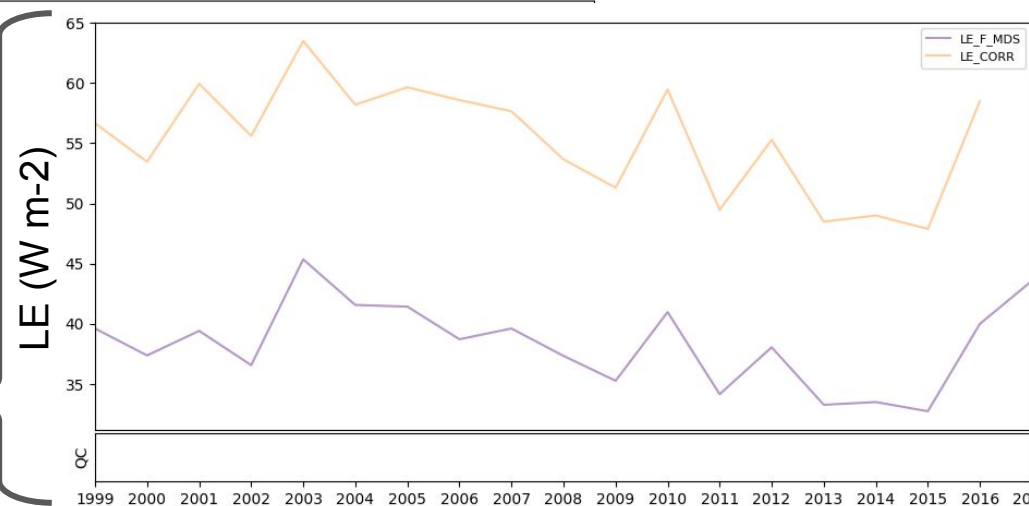
IT-Lav



LE (W m⁻²), hourly and yearly

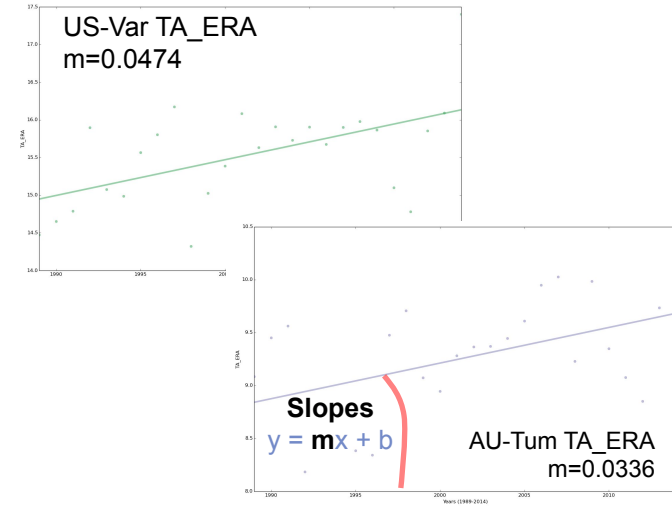


US-MMS



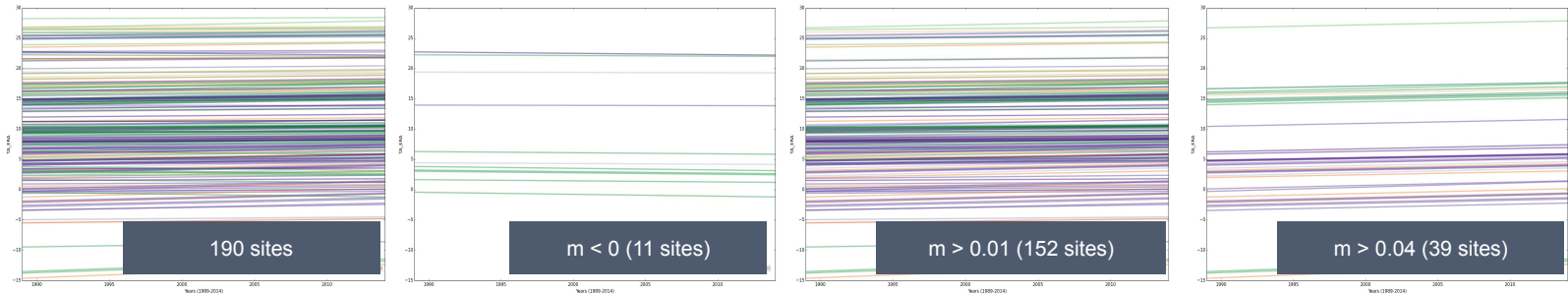
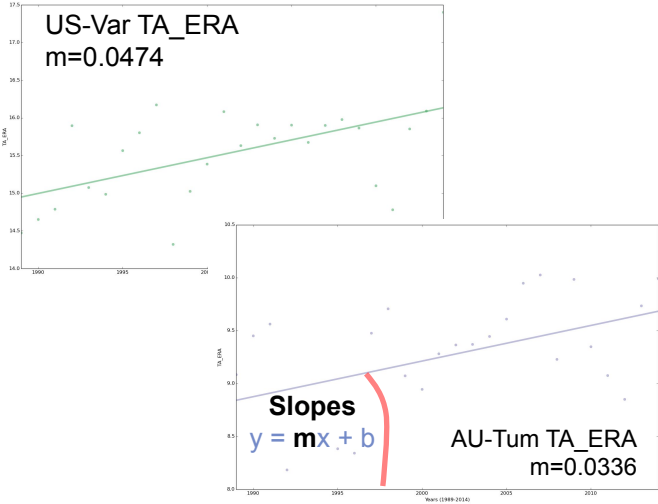
ERA Downscaling

(ERA-I) 1989-2019 → (ERA5) 1979-Now



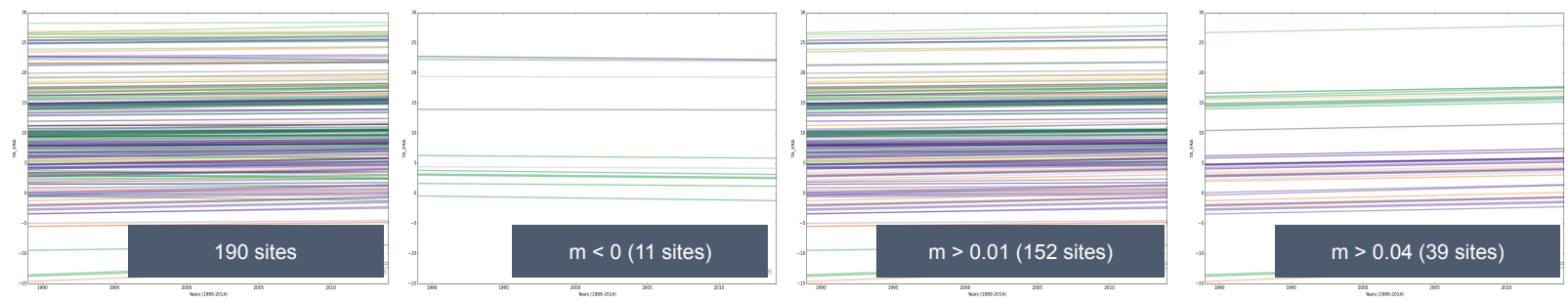
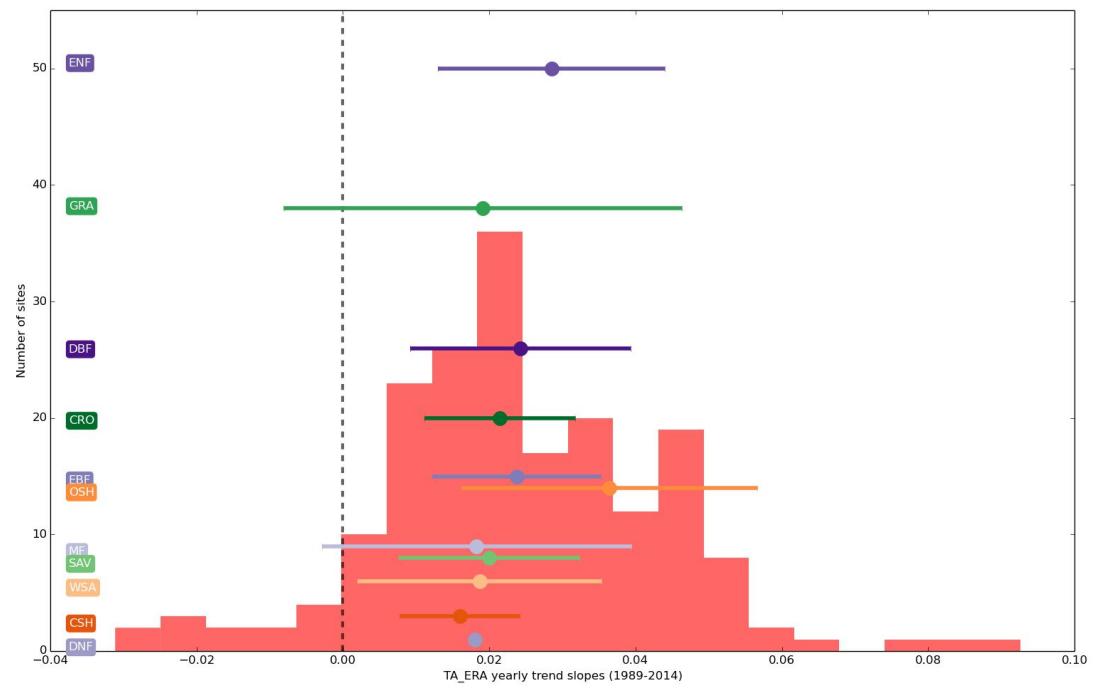
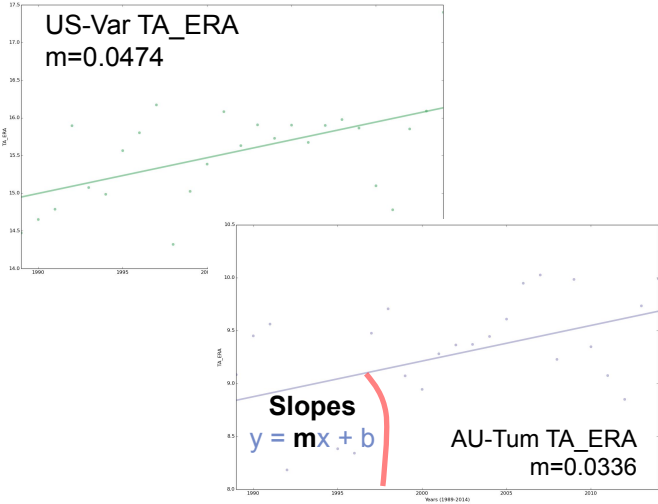
ERA Downscaling

(ERA-I) 1989-2019→(ERA5) 1979-Now



ERA Downscaling

(ERA-I) 1989-2019→(ERA5) 1979-Now



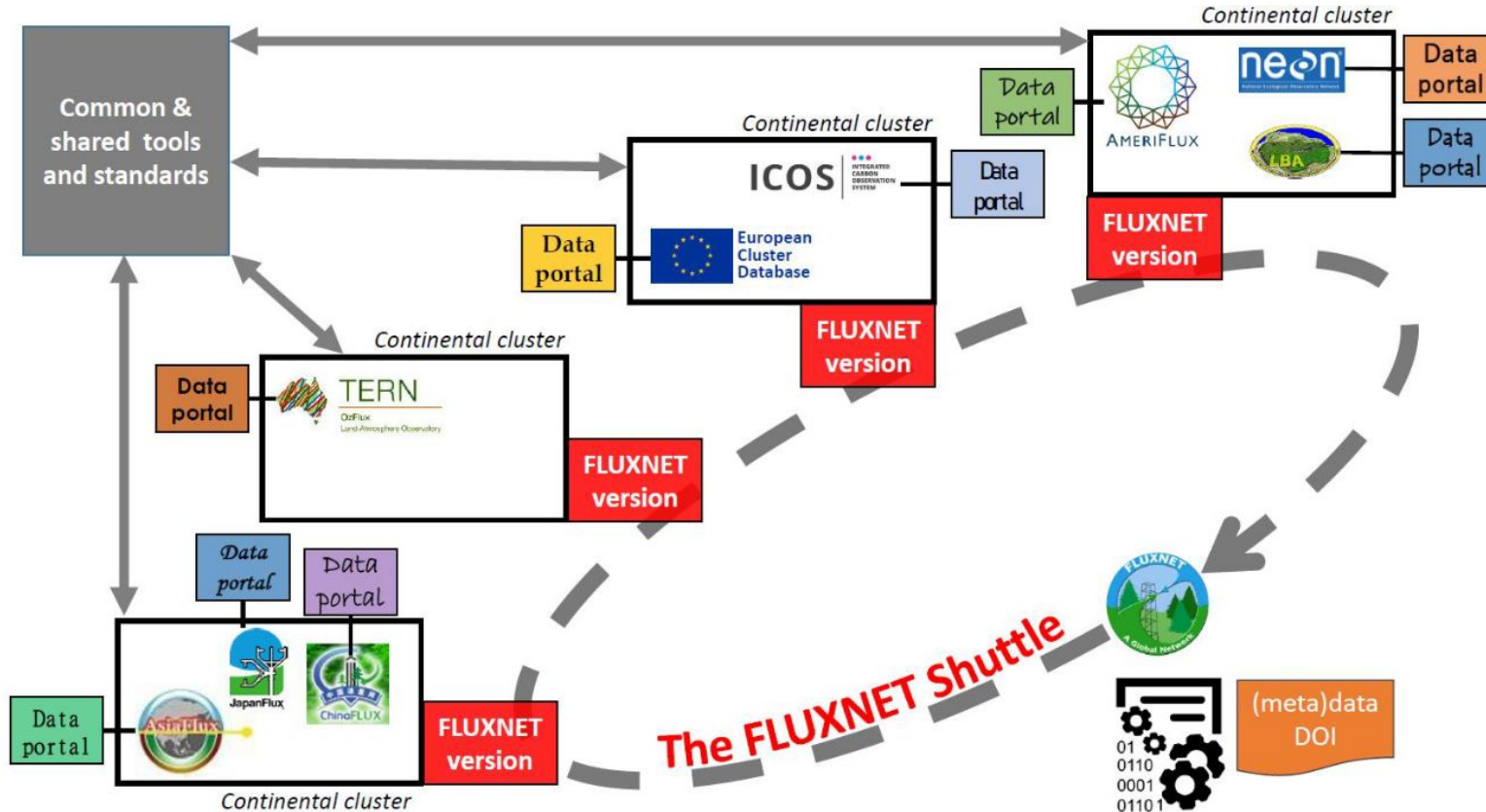
Agenda

- ONEFlux overview
- Workflow of a ONEFlux run
- Metadata collection for ONEFlux processing
- Data QA/QC for AmeriFlux BASE data and impacts on ONEFlux runs
- FLUXNET data products created with ONEFlux
- Future of FLUXNET regional and global data products



*Dario
Papale*

Future of FLUXNET regional and global data products



Q&A

Thank you!

Questions?



Reference paper:

Pastorello, G., Trotta, C., Canfora, E. *et al.* (287 co-authors). ***The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data.***

Scientific Data **7**, 225 (2020).

<https://doi.org/10.1038/s41597-020-0534-3>

Code available:

<https://github.com/FLUXNET/ONEFlux>

Contact us: ameriflux-support@lbl.gov